

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 09:44:53 ; Search time 86.21 Seconds
(without alignments)
4590.135 Million cell updates/sec

Title: US-09-817-538-2
Perfect score: 1611
Sequence: 1 atgtctgggtctctgcgcg.....tccctcagtttcttcccc 1611

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1498	93.0	2111	1	US-08-528-255A-2
2	1498	93.0	2111	1	US-08-717-365-2
3	502.6	31.2	1943	4	US-09-282-305-7
4	501.8	31.1	2019	4	US-09-282-305-5
5	460.8	28.6	1576	4	US-09-282-305-9
6	459	28.5	1826	4	US-09-282-305-1
7	392.4	24.4	600	4	US-09-385-982-402
8	329.2	20.4	1475	4	US-09-282-305-3
9	204.6	12.7	841	4	US-08-998-416-317
10	77.2	4.8	7218	1	US-08-232-463-14
11	50	3.1	3489	2	US-08-728-323A-1
12	50	3.1	32207	2	US-08-770-379-20
13	50	3.1	32207	4	US-08-757-669A-20
14	50	3.1	32207	4	US-09-230-371A-20
15	48.4	3.0	390	4	US-09-197-649-7
16	44.2	2.7	1931	2	US-09-130-114-2
17	43.8	2.7	3223	1	US-07-980-528-1
18	43.8	2.7	6755	3	US-08-931-999-4
19	42	2.6	480	2	US-08-726-306A-185
20	41.8	2.6	5433	3	US-08-929-329-1
21	41.6	2.6	1042	4	US-09-276-531-77
22	41.6	2.6	43795	3	US-08-742-185-101
23	41.2	2.6	2233	1	US-08-145-705A-1
24	40.8	2.5	49136	4	US-09-422-869-1
25	40.8	2.5	51259	3	US-08-781-891-209
26	40.6	2.5	289	4	US-09-007-005-17
27	40.6	2.5	289	4	US-09-244-796-17

28	39.6	2.5	15378	3	US-08-785-420-1	Sequence 1, Appli
c 29	39.2	2.4	16442	3	US-08-781-891-208	Sequence 208, App
30	38.4	2.4	2277	1	US-08-676-967-2	Sequence 2, Appli
31	38.4	2.4	2277	1	US-08-676-974-2	Sequence 2, Appli
32	38.4	2.4	2277	2	US-09-098-487-2	Sequence 2, Appli
33	38.2	2.4	2223	1	US-08-557-073-4	Sequence 4, Appli
34	38	2.4	1298	3	US-08-948-705-3	Sequence 3, Appli
c 35	37.8	2.3	2188	1	US-07-865-662F-10	Sequence 10, Appl
c 36	37.8	2.3	2188	4	US-08-374-219B-10	Sequence 10, Appl
37	37.6	2.3	43676	3	US-09-356-952-12	Sequence 12, Appl
38	37.6	2.3	80246	4	US-09-078-294-4	Sequence 4, Appli
39	37.6	2.3	80595	4	US-09-078-294-3	Sequence 3, Appli
40	37.4	2.3	301	2	US-08-332-766A-23	Sequence 23, Appl
41	36.8	2.3	3211	2	US-08-574-959A-8	Sequence 8, Appli
42	36.8	2.3	3211	4	US-09-357-014-8	Sequence 8, Appli
43	36.8	2.3	3901	2	US-08-574-959A-6	Sequence 6, Appli
44	36.8	2.3	3901	4	US-09-357-014-6	Sequence 6, Appli
45	36.6	2.3	1136	1	US-08-186-981-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-528-255A-2
Sequence 2, Application US/08528255A
Patent No 5659016
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
TITLE OF INVENTION: FURUKAWA, YOICHI
TITLE OF INVENTION: RPDL PROTEIN AND DNA
ENCODING THE SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANTIS, P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/528,255A
FILING DATE: September 14, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP6-227876
FILING DATE: 22-SEPTEMBER-1994
APPLICATION NUMBER: JP7-183763
FILING DATE: 20-JULY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teriyence F. Chapman
REGISTRATION NUMBER: 32549
REFERENCE/DOCKET NUMBER: Furuya Case 1335
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2111
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal lung cDNA library
FEATURE:

```

; NAME/KEY: CDS
; LOCATION: 64..1512
; IDENTIFICATION METHOD: experimental examination
US-08-528-255A-2

```

Query Match	93.08%	Score 1498;	DB 1;	Length 2111;
Best Local Similarity	99.38%	Pred. No. 0;		
Matches 1504; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;

QY	97	gggaggcgagcaagaatggcgccagagcgcaagggcaccggaggaaagtcgtgttactacg	156
Db	50	GGGAGGCGAGCAAGATGGCGCAGACCGCAGGGCACCCGGAGGAAGCTCTGTACTACTACG	109
QY	157	acgggagtggtgnaattactattatygacaagggcccaatlyaagccctcaccgaatcc	216
Db	110	ACGGGATGTTGGAATTAATAATTATGACCAAGGCCCAATGAAGCCTCACCGAATCC	169
QY	217	gcatactcataatttgcctgtcctaactatggtctctacggaaaaatgynaatctatcgcc	276
Db	170	GCATGACTCATAAATTGCTGCTCAACTATGGTCTCTACCGAAAAATGGAATCTATCGCC	229
QY	277	ctcacaagccaatgctgaggagatgacccaagttaccaagcgatgactacatlaattct	336
Db	230	CTCACAAAGCCAATGCTGAGGAGATGACCAGTACCACAGCGATGACTACATTTAAATTC	289
QY	337	tgcgctccatccgtccagataaacaatgctcgagttacagcaagcagatgacagatatcaacg	396
Db	290	TGGCGCTCCATCCGTCAGATAACATGTCGGAGTACAGCAGACAGATGCAGAGATTCAACG	349
QY	397	ttgttgaggactgtccagtatcgaatggcctgttgaattctgtcagttgtctactgtg	456
Db	350	TTGGTGAGGACTGTCCAGTATTGCATGGCCGTGTTGAGTTCTGTCAAGTTGCTACTCGTG	409
QY	457	gtctctggccaagtgtctgtgaaacttaataagcagcagacagacatcgccgtgaattgg	516
Db	410	GTTCTGTGGCAAGTCTGTGAAACTTAATAAGCAGACGACGGACATCGCTGTGAATTGGG	469
QY	517	ctgggggacctgcaccatgcaagaagtcaggagcatctgctctgttacctcaatgata	576
Db	470	CTGGGGCCTGCACCATGCAAAAGAGTCCGAGGCATCGGCTTCTGTACGTCATGATA	529
QY	577	tcgtcttgccatccctgggaactgtctaaagtataccaccagagggtgctgtacattgacatg	636
Db	530	TCGTCCTGGCCATCCTGGAACTGCTAAAGTATCACCAGAGGGGTGCTGTACATTGACATTG	589
QY	637	atatcaccaatggtgacggcggtggaagagcccttcaaccaacggaaccgggtcatgactg	696
Db	590	ATATTCAACCATGGTGAACGGCGTGAAGAGGCCCTTCTACACCACGGAGCCGGGTCAATGACTG	649
QY	697	tgctccttcataaagtatgagagtagtacttcccaggaactggyggacctacgagataccggg	756
Db	650	TGTCCTTTCATAAGTATGAGAGTACTTCCCAGGAACCTGGGGACCTACGGGATATCGGGG	709
QY	757	ctggcaaaagacaagtataatgctgtttaactacccgcctccgagacggygattgatacgaagt	816
Db	710	CTGGCAAAAGGCAAGTATTATGCTGTTAACATACCCGCTCCGAGACGGGATTGATGACGAGT	769
QY	817	ccatagagccatttccaagccggtcatgtgccaaagtaatgagatgttccagccatgtg	876
Db	770	CCTATGAGGCCATTTTCAAGCCCGGTCAATGCCAAAGTAATGAGATGTTCCAGCCCTAGTG	829
QY	877	cggtggtctacagtggtggtcagactcccatctctggggatcggttagtgttcaatc	936
Db	830	CGGTGCTCTACAGTGTGGCTCAGACTCCCTATCTGGGGATCGGTTAGTGTCTCAATC	889
QY	937	tatcatcaaaagacaacggccaagtgtgtgynaatttgtcaagagctttaacctgtcctatgc	996
Db	890	TAACATATCAAAAGSACACGCCCAAGTGTTGGAATTTGTCAAGAGCTTTAACCTGCCATATGC	949
QY	997	tgatgtcgggagcggtgtgttacaaccatcgtaacgttgcgggtgtcgtgacataatgaga	1056
Db	950	TGATGCTGGGAGCGGTTGGTTACACCATTCGTAACGTTGGCCCGGTGCAGAGACATATGAGA	1009

OY	1057	cagctgtgccccttgatcgcgagatccctaatagaagcttcacaaatgatgacttctgaaat 		1116
Db	1010	CAGCTGTGCCCCCTGGATCGAGAGATCCCTAATGAGCTTCCATTACATGACTACTTTGAAT 		1069
OY	1117	actttgaccagatttcaagtccacatcacgtccctccaatatgactaacccagaacacga 		1176
Db	1070	ACTTTGACCAGATTTCAGACTCCACATCATCAGTCCCTTCCAATATGACTAACCAAGAACA 		1129
OY	1177	atgagttacctggagaagaatccaacaacgcgactgttttagaaaccttagaatgtctgccgcacg 		1236
Db	1130	ATGAGTACTCTGGAGAAGATCAACACAGCGACTGTTTGAGAACTTAGAATGCTGCCGCACG 		1189
OY	1237	caacctgygtccaacacgcgaggcgalttccttgagacgcacatccctyagagagatgycgatg 		1296
Db	1190	CACCTGGGGTCCAAATGCAGGCGATTCCTTGAGGACGCCATCCCTGAGGAGAGTGGCGATG 		1249
OY	1297	aggacgaagaacgacccctgacaacgcgatctcgatctgtctccttgacaaacyaatgtgcct 		1356
Db	1250	AGGACGAAGACGACCCCTGACAAAGCCGATCTCGATCTGCTCCTTGACAAACGAATTGCCCT 		1309
OY	1357	gtgaggaagaagtcttcgcgattcttgaaagagagggaggggycgcgaagaactcttcca 		1416
Db	1310	GTCGAGGAAGAGTTCTCCGATTCTGAAGAGGAGGGAGAGGGGGCCGCAAGAACTTTCCA 		1369
OY	1417	acttcaaaaaagccaagaagatcaaaaacagaagatbaaaaaagagaagaagaccagaagaga 		1476
Db	1370	ACTTCAAAAAAGCCAAAGAGAGTCAAACACAGAGGATGAAAAAGAGAAAAGACCACAGAGAGA 		1429
OY	1477	agaaaagaaatcacccgaagagagagaaaaccaagagagagaaagccagaagccaaagyggtca 		1536
Db	1430	AGAAAGAAGTCAACCGAAGAGAGAGAAAAACCAAGAGAGAGAAAGCCAGAAGCCAAAGGGGTCA 		1489
OY	1537	aggagagagccaagtgtgcctgtaatgtagacctctcaagctctgtgcttcctgtgtgattcccc 		1596
Db	1490	AGGAGGAGGTCAAGTGTGGCTGATGACCTCTCCAGCTCTGGCTTCTGTAATCCCCT 		1549
OY	1597	caagtttcttcccc 1610 		
Db	1550	CACGTTTCTTCCCC 1563 		

RESULT 2
 US-08-717-365-2
 Sequence 2, Application US/08717365
 Patent No. 5763182
 GENERAL INFORMATION:
 APPLICANT: NAKAMURA, YUSUKE
 APPLICANT: : FURUKAWA, YOICHI
 TITLE OF INVENTION: RPDL PROTEIN AND DNA
 TITLE OF INVENTION: ENCODING THE SAME
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FLINN, THEIL, BOUTELL & TANIS, P.C.
 STREET: 2026 Rambling Road
 CITY: Kalamazoo
 STATE: Michigan
 COUNTRY: USA
 ZIP: 49008-1699
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
 COMPUTER: IBM PC/XT/AT Compatible
 OPERATING SYSTEM: MS-DOS 5.0
 SOFTWARE: Wordperfect 5.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/717,365
 FILING DATE: 23-SEP-1996
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/528 255
 FILING DATE: September 14, 1995
 APPLICATION NUMBER: JP6-227876
 FILING DATE: 22-SEPTEMBER-1994

; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)..(1610)
US-09-282-305-7

Query Match 31.2%; Score 502.6; DB 4; Length 1943;
Best Local Similarity 65.5%; Pred. No. 2.6e-143;
Matches 768; Conservative 0; Mismatches 399; Indels 6; Gaps 2;

QY 134 gaggaagctgttactactacgacgaggatgttgaataattactattatgacaaaggcca 193
Db 122 gcgcgcgtgtgtactcttctacccggatgttggcaactactactacgagcgagcca 181
QY 194 cccaatgaagcctcacggaatccgcatgactcataattgtctgaactatgtctcta 253
Db 182 tccgatgaagcgcgcacgcataccgatgacgacactcgtctgtgcgcgtacgacctct 241
QY 254 ccgaaaatggaatctatcgccctcacaaagccaatgctgagagatgaccaaagtaacca 313
Db 242 caaccagatgcaggtgtacccgcccaaccggccgcgcgacccgacctctgccccttcca 301
QY 314 cagcgatgactacattcaattcttgcgctccatccgtccagataacatgtcggagtaacag 373
Db 302 cgcgcagcactacatacacttctcgtcgcctcgtcgcgcgcgaaacgcagcagcagccagat 361
QY 374 caagcagatgcagagatccaacgctgtgtaggagactgtccagtaattcgatggcccttga 433
Db 362 ccgcctgtcgaagcgttcaacgctgcgcgagagactgtcccgcttctgcagcgccctacag 421
QY 434 gtctgtcagttgtctactgtgtctgtgtgcgaagtgttgaaccttaataagcaagca 493
Db 422 ctctgcagaacctatgcgggcgcctcgtcgcggggcggtlcaagctcaaccaacggcca 481
QY 494 gacggacatgcgcgtgaattggcggtggggcctcaccaatgcaaaagaagtcggaagcatc 553
Db 482 ---tgacatgcacaactgctcggggggtgcgcacgaacgaagaagtgcgagggcgtc 538
QY 554 tggcttctgttacgtcaatgatalcgtcttggccactcctggaactgtctaagatcacca 613
Db 539 gggcttctgtctacgtcaatgacatcgtgtcgcacatactcgagctgtcgaagcatcacga 598
QY 614 gagggtgctgtacattgacattgatatccacatgtgacggcgttgaagaggccttcta 673
Db 599 gagagttctgtatgtcgatatcgataccaccatgtgtgagtgagagggcttcta 658
QY 674 caccacgacccggtcactgactgtgtccttccataagtatgagagtaacttcccaaggaa 733
Db 659 cacaacagatagggttatgactgtctcgttccacaagtttgttattatttcccaaggaa 718
QY 734 tggggacctacgggataccgggggtgcaaaagaattatgtgttlaactcccgct 793
Db 719 aggggatatccgtgacattgggcaactcaaaaggagtaactaccctgaatgtccctct 778
QY 794 ccgagacggatgtatgacgagtcctatgagggcaatttcaagccggtcatgtccaagt 853
Db 779 agatgatggatgtatgataaagctacacagtcctttaaagccaatcatgtggcaagt 838
QY 854 aatggagatgttccagcctagtgcggtgttcttacagtggtgcacactccctatctg 913
Db 839 tatggaggtttccgcctcgtgtgcagttgtgtcttcaagttgtgtcattccttctgtg 898
QY 914 ggatcggttaagttgttccaatctatctatccaagaagcaacgcgaagtgtgtgaatttgt 973

Db 899 ggataagttggcgtgtcttcaacctctcaactcaataaaggctacgtcggaatgtgttaggtatat 958
QY 974 caagagcttcaacctgacctatgtctgatgtcgtgagggcggtgttacaccattcgtaacgt 1033
Db 959 gaggcttcaacglttccattgttctccttgggtgtgtgtgatataccaataagaatgt 1018
QY 1034 tggccggtgtgacataatgacagactgtgcccctgataacgagatccctaagt 1093
Db 1019 tgcacgctgtgtgtatgagactgtgactgtcttggccaagagcctgaagaagaat 1078
QY 1094 tccatacaatgactacttgaataacttggaccagatttcaagctccacatcagtcctc 1153
Db 1079 gcctgttaatgactataatgataactcgtgtccagattacactcttcaatgtgcaccaag 1138
QY 1154 caatatgactaacgaaacacgaatgagtacctgagagaatcaaacagcagactgttga 1213
Db 1139 taacatgagaaacaaataacacgcacacacactggatgatatacagatctaaacttctga 1198
QY 1214 gaaccttagaatgtgtccgcacgcacaccttgggttccaaacgcagcgca---ttcctgagga 1270
Db 1199 taacttcaaaacttcgacatgtctcctagtgtccacttcaagagagaggttccctgacac 1258
QY 1271 cgccatccctgaggaagtggtgcgatgagca 1303
Db 1259 agaataacctgagcaagatgaagatcaagatga 1291

RESULT 4

US-09-282-305-5
; Sequence 5, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (140)..(1459)
US-09-282-305-5

Query Match 31.1%; Score 501.8; DB 4; Length 2019;
Best Local Similarity 65.3%; Pred. No. 4.7e-143;
Matches 753; Conservative 0; Mismatches 397; Indels 3; Gaps 1;

QY 121 cgcagggcaccggggaagagctgttactactacgacgggatgttgaattactatt 180
Db 189 cgaaggtgtcgaagggcgcgtctgtctacttcaacgacggaggtggcaactactact 248
QY 181 atgacaagggccaccaatgaagcctcacccgaatccgcatgactcataattgtctca 240
Db 249 acgggcagggccaccgatgaagccgcacgcgcatcgatgaccacagcgctgcggcc 308
QY 241 actatgtctctacgaaaaatggaatctatcgccctcacaaagccaatgtgagaga 300
Db 309 gctacggcctcctcgaccagatgcaagtgttccgacctcacacctgtccgcgacccgac 368
QY 301 tgaccaagtaccacagcgatgactacatlaattcttgcgtccatccgtccagataaca 360
Db 369 tctgcgcttccacgcgcgacgattacgtctccttctcgtcgtccgtacccccgaagc 428

QY	361	tgctcgagtagtacgacgaacgacgatgacagatccaacgtgtgtgaagcactgtccagttatcg	420
Db	429	agcaggaaccagatccgcgcgtccaagcgcctccaacgctcgcgcgagagactgtccccccttcgc	488
QY	421	atgacgtgtttgagttctgtcagttgttctactgtgtgttcctgtgcgaagtgtgtgaac	480
Db	489	acgctctctacagtttctgtcagacgtacgcgggggctcgtgttgcgcgccgtcaagc	548
QY	481	ttaataagcagcagacgcgacatcgccgtgaatttggcgtcggggccctgcaccatgtcaaga	540
Db	549	tcaaccatgtgcc--tgaatacgcacatcaactgtggccgcgcgactccaaccagccaaga	605
QY	541	agtcgcagacatctgtgcttctgtttaacgtcaatgatatalcgttctttggccactctgaactgc	600
Db	606	agtcgcagacgtccgcggtttgtctatgtttaatgatcatgttctctgcacatcctcgaagctcc	665
QY	601	taaaagtataccacagaggtgtgtgtacattgacatltgatattccaatgtgtgaacgcgtgcg	660
Db	666	tcaagtaaccacacagcgcgttctgttacctgtgacatltgatattccaacagggagacgcgtgcg	725
QY	661	aagaagcctctacacacacagccgggtcatcagctgtgtccttccataaagtatgagagat	720
Db	726	aggaagcctttatataccacagacccgggtgatgtacagcttccatccacaagtttgagatc	785
QY	721	acttccacgaagactggggacctaacgggataccggggtcgtgcgaagaagatlatgtctg	780
Db	786	attccctgtgacagaggggacatctcgtatgttggcaccatcaaggtaaatattactccc	845
QY	781	ttaactacccgcctccgagacgggatttgtatgacgaagctctatgagggccatttccaagccgg	840
Db	846	tgaatgttccctgtgacgcgatgtatttgtatgtagagactaaccagctcgttgttccaagccaa	905
QY	841	tcatgtccaaagtaatgagagatgttccagcctagtgcgtgtgtcttacagttgtgcctcag	900
Db	906	taatgggcgaagtgatgtgaggtcttcaaccctgtgtcagtcgtctccaagttgtgtgcgg	965
QY	901	actccctatctggggatcgggttaagttgtcttcaatctatctataccaagaagacagccaagt	960
Db	966	attcatgtgcggtgacagagttggcgtgttccaacctctctaataaggtcaacgcagaat	1025
QY	961	gtgtggaattgttcaagagctttaaacctgcctatgtctgatatcgtgcggagcggtgttaca	1020
Db	1026	gtgtaagatttatgaggtcccttcaacgctcccgctgttgtctgtcttgtgtgtgtgtata	1085
QY	1021	ccattcgtlaacgttgcgcgggtgctgacatatgagacagcgtgtggccctgatacggaga	1080
Db	1086	ccataagaacgtytgacgcgtgtgtgtgtctacgagacagagagttggccctgtgtcatgagc	1145
QY	1081	tcacctaatgagcttccaataatgactactcttgaataactttgaccagatttcaagctcc	1140
Db	1146	tcactgacaagatgtccaccctaaatgagtactatgagtattttgttccagattcacacttac	1205
QY	1141	acatcagtccttccaatatgactaataaccagaaacagaaatgagtactctggagaagaatcaaac	1200
Db	1206	atgtcgtctccaagtatacatgtgagaaataaaaaacacacacgcgcatcaattggtatgacataaat	1265
QY	1201	agcgactgttttgagaacacttagaatgctgcgcgacgcacactcgtgggttccaaacgcagcgca	1260
Db	1266	caaaactctagataatcttccaanaactccgacatgtctcctagtgttcaagtagc	1325
QY	1261	ttcctgaaggagc	1273
Db	1326	gacctcctgaagc	1338

RESULT 5
US-09-282-305-9
; Sequence 9, Application US/09282305
; Patent NO. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.

```

; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1576
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (38)..(1336)
US-09-282-305-9

```

Query Match	28.6%;	Score 460.8;	DB 4;	Length 1576;
Best Local Similarity	62.0%;	Pred. No. 1.3e-130;		
Matches 764; Conservative	0;	Mismatches 462;	Indels 6;	Gaps 2;

QY	159	gggagtggtggaattactatctatgtgacaaggccacccaatgaagcctcacggaatccgc	218
Db	80	ggagatggttgcgaatgtctacttttgggccaaatcaccccataagccacatcgctctcgt	139
QY	219	atgactcataatttgtctgtccaactatggtctctaccgaaaaatgyaatctatcgccct	278
Db	140	atgacacatcaccttgttcttcttcatalgagcttcatcaaaaagtggagatatataaggcca	199
QY	279	cacaaagccaatgctgagagatgaccaagtaccacagcgaaagtacatatcaattcttg	338
Db	200	cacaaagcatatccaataagagcttgccccaatlccatctgtgatattaigtggaattcttg	259
QY	339	cgtccatccgtccagataaacatgtcggagtaacagcaagcagatgcagagatlcaaagtt	398
Db	260	caccygataactcctgattccaccagccctatatgcaagtgaaactaactagatacaactct	319
QY	399	ggtgagactgtccagtalctcgatgscgtlttgagttctgtcagttgtctactgtgtg	458
Db	320	ggagaagactgtccggtctttagataaattgttgagttctgcacaatctatgcyggsgga	379
QY	459	tctgtgccaagtctgtgaaccttaataagcagcagacgcagacatcgccgtgaattggct	518
Db	380	aacttagatgcgtctgcagattaalcatcaataaatatgtgacattgccaattaattggct	439
QY	519	gggggcctgcaccatgcaagaagtcaggagcatctggtctgtttaagtcgaatgataac	578
Db	440	ggtaggtacatcatcatgccaanaagtgtgaggtctcaggtctctgttaacattaatgatcta	499
QY	579	gtcttgccatcctgtgaactgtctaagaatatcaccagaggtgtgttacattgacattgat	638
Db	500	gtattaggaattctgtgagcttctcaagtaaccatgccaaggttcttataattgacattgat	559
QY	639	attcaccaatgtgacgycgtggaagggcctctacacacagacccggtcatgactgtg	698
Db	560	glccatcatggagatgtgagltggaagagcccttlattcacatgcacaggtaatgactgtg	619
QY	699	tccttcataaagtatygaga--gtacttccaggaactgggacctaagcatgataccgg	755
Db	620	agttccacaagaatatgtgacctgttcttccgtgaacaggtgatattaagagatatagga	679
QY	756	gctgcaaaagacaagtatatgtctgttaactaccgcgtccggagacgggattgatgacgag	815
Db	680	gaaagggaaagaaatatlatgtctatcaacatctccacttaagaatggtgatatgacact	739
QY	816	tcctatgagccaatttcaagccggtcatgtccaagttaatggagatgttccagcctagt	875
Db	740	agcttactcggcctttaaacaacaatctatggccaagttgttgagacatatctgcgcgtg	799
QY	876	gcggtggtcttcaagtgtgtgctcagactccctatctggggatcgtgttaggtgtctccaat	935

Db 800 gctattgttcttcaatgtggygctgattcattgcgagggatcgtttagcgcttccaat 859
QY 936 ctatctatcaaaagacacgccaagtgtgtgaattgtcaagagctttaacctgctatg 995
Db 860 ctctctatgaagccatgctgtaagtgtgaagttgttcaagaanaattcaatattccctt 919
QY 996 ctgattgtggaagcggtgtgttacaccattctgaacgttgcgcgtgtgtgacatatgag 1055
Db 920 ctgtgaactggaagtgtgtgtatcaccaaggaagtgtagcacggtgttggcgctgtgaa 979
QY 1056 acagctgtgcccctgatacgagatccctaattgagcttccatacatagactactttaa 1115
Db 980 actgggtctcttttagacacagaactcccaaatgagattccaaaataatgaatatattgag 1039
QY 1116 tactttgaccagatttcaagctccacatcagtccttccaatatgtactaaccaagacag 1175
Db 1040 tactttgctccagattatcatttgaagttccaatttgaacatgacacatttgaacagt 1099
QY 1176 aatgagtacctggaagaagatcaaacacagcagctgtttgagaaccttagaatgtgcgcgac 1235
Db 1100 aagacctatctcagttccaatccaagtgcaagtgtgagaaggtttgcgtacatacagcat 1159
QY 1236 gcaacctgggtccaaacagcgagcttcttgagagcgccatccctgagagagtgcgcat 1295
Db 1160 gctcctgtgttcaaaagcaagaggttctctcccgattttatatcccgacttt---gat 1216
QY 1296 gaggaacgaagacacccctgaacagcgcatctcgatctctcctctgacaaacgaattgcc 1355
Db 1217 gaagatgaattgatacctgatacgatgtgtgacacagcacactcaagacaagcagattcac 1276
QY 1356 tgtgagaagaagttctccgattctgaagaga 1387
Db 1277 cgtgatgatgatactatgaaggtgacaatga 1308

RESULT 6
US-09-282-305-1
; Sequence 1, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1826
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (29)..(1405)
US-09-282-305-1

Query Match 28.5%; Score 459; DB 4; Length 1826;
Best Local Similarity 61.8%; Pred. No. 5.1e-130;
Matches 729; Conservative 0; Mismatches 450; Indels 0; Gaps 0;
QY 131 ccgagagaaagtctgttactactagacggygagtgtgaaattactattgagacaagg 190
Db 88 ccgccgcgcgtcagctatttctacgagccgtgcgactcgagactactactagggcaagg 147
QY 191 ccaccaatgaagcctcacccgaatccgcatgataataattgtgtcctaactatgtct 250
Db 148 tcacccgatgaagcccaaccgataccggaatgagcgactcgtgtgttccactacgacct 207

QY 251 ctaccgaaaatatgaaatctatcgccctcaacaagccaatgctgagagatgaccaagta 310
Db 208 ccaccgcctctcgcagctctcccgccctaccgcgcctctlgagccgacatccgcgcct 267
QY 311 ccacagcatgatacatattaattcttgcgtccatccgcgtccagataacatgtcagta 370
Db 268 ccactccgacgactacgtcgtcttctcctgcgtcccgccacccgaaaccgggtgtcctga 327
QY 371 cagcaagcagatgcagagattcaacgttgtgtgagactgtccagtatccgactgtgt 430
Db 328 cccgcgcgcattaaagcgtttaacgtcgtgtgagactgcccggtgttcgacggtctct 387
QY 431 tgagttctgtcagttgtctactgtgtgttctgtggaagtgtgtgaacttaataagca 490
Db 388 ccccttctgcagcgctccgctgtggggaagcattcgcgcgcgtccaagcttaaccgcg 447
QY 491 gcagacgacatcgccgtgaattgggctggggcctgcaccatgcaagaagtcagggc 550
Db 448 ggaacgcgacatcacccgtcaacttggcgggcgccctccacacgcgaagaagcgaggc 507
QY 551 atctgctctgttcaatgataatgatcgtcttggccatccctggaactgttaagatca 610
Db 508 ctccgggtctgtctacgtlcaacgacatcgtctcgcctccatcctcgagctcctcaagttcca 567
QY 611 ccagaggtgtgttacaattgacattgatatcccaatggtgacgcggtggaagggcct 670
Db 568 cagcggtgtgtatagtacatattgatgtgccacatgagatggtgagggagggcct 627
QY 671 ctacaccacgagccggtcatgactgtgtccttcataaagatagagagacttccag 730
Db 628 ctactacaacccgagtcagtactgttcccttcacaagatagggatctttccctcg 687
QY 731 aactgggacctaagtgatataccgggctgtgcaagaacgaattatgtctgttaactacc 790
Db 688 tactgacatatcactgacgcttggggcagccgaagggaagcattatgtcttgaatgttcc 747
QY 791 gctccgagacgggattgtatgacgagctcctatgagggcatttccaagccggtcatgtccaa 850
Db 748 cctgagtatggtatcgtatgacacacaccccttctgtgtctgtttcaatgcatcattaagaa 807
QY 851 agtaatgagatgttccagccctagtgctgtgtcttacaagtgtgtcagactccctatc 910
Db 808 agttatgaggtttatcagccagacgctgtgttctcctccaatgctgagctcgtactcttggc 867
QY 911 tgggagtcggttaggtgttcaatctatctatcaaaaggaacgacgccaagtgtgtgaatt 970
Db 868 tggagacaggttaggtgttcaacctgtctgtgaagggtcatgtcgtactgtcctcgltt 927
QY 971 tgtcaagagcttcaacctgcctatgtctgtagtgggagcggtgttaaccattcgttaa 1030
Db 928 ccttagtctgatacatgttctctatgatgtgtttaggtgtgtgaggttacaaccatcagaaa 987
QY 1031 cgtgcccgtgtgtgacatatgagacagctgtggccctlgatagcgaagatccctaataga 1090
Db 988 tgttgacgctgtcgtgtgtctacgagaccgcagctgtgtgtgagttgaaactgtatacaa 1047
QY 1091 gcttcatacaatgactactttgaataactttgaccagagatttcaagctccacatcagtc 1150
Db 1048 gctgccttacaatgatactatgagtacttggccctgattatatacttcatataccaacc 1107
QY 1151 ttccaatagactaacacgaacagaaatgagtaacctgtgagaagaatcaaacagcgactgt 1210
Db 1108 aaaaagtgtgaaaacctgatacaccaaaagacttggagaacataaagaacatgatatt 1167
QY 1211 tgagaaccttagaatgtgtccgcagcgcactgtgggttccaacgcagcgatctcctgagga 1270
Db 1168 ggagaacctgttcaaaagatagaacatgttcccaagcactcaattccatgacagaccgtcaga 1227
QY 1271 cgcacatccctgagagagtggtgcgatgagacgaagaagca 1309
Db 1228 ccctgaagctccagagaggagaagagagacatgagcaaa 1266

```
RESULT 7
US-09-385-982-402
; Sequence 402, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 402
LENGTH: 600
TYPE: DNA
ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(600)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-402

Query Match          24.4%; Score 392.4; DB 4; Length 600;
Best Local Similarity 98.1%; Pred. No. 5.5e-110;
Matches 407; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1183 acctggaagaatcaaacagcagctgttgagaaccttagaatgtccgcagcacctg 1242
|||||
Db 1 acctggagaagatcaaacagcagcgtgttgagaaccttagaatgtccgcagcacctg 60
|||||

QY 1243 ggggtccaaacgagcgatcttcctggagcgccatcccttgagagagtg-gcatgagac 1301
|||||
Db 61 ggggtccaaatgcagcgatcttcctggagcgccatcccttgagagagtg-gcatgagac 120
|||||

QY 1302 gaagacgaccttgacaagcgcatctcgtatctgtctccttgacaacgaaattgctgtgag 1361
|||||
Db 121 gaagacgaccttgacaagcgcatctcgtatctgtctccttgacaacgaaattgctgtgag 180
|||||

QY 1362 gaagagttctccgattctgaagagagggagagggggcgccgcaagaactctccaacttc 1421
|||||
Db 181 gaagagttctccgattctgaagagagggagagggggcgccgcaagaactctccaacttc 240
|||||

QY 1422 aaaaaagccaagagagtcataaaacagagatgaaaaagagaagaccagagagagaagaaa 1481
|||||
Db 241 aaaaaagccaagagagtcataaaacagagatgaaaaagagaagaccagagagagaagaaa 300
|||||

QY 1482 ggaatcacccgaagagagagaacccaagagagagagccagaagccaagggtcaagagag 1541
|||
Db 301 ggaatcacccgaagagagagaacccaagagagagagccagaagccaagggtcaagagag 360
|||

QY 1542 gaggccaaagtggcctgaatgagccctccagctctggtctgctgagtcacct 1596
|||||
Db 361 gaggtcaagtggcctgaatgagccctccagctctggtctgctgagtcacct 415
|||||

RESULT 8
US-09-282-305-3
; Sequence 3, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
```

```
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 3
LENGTH: 1475
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
; NAME/KEY: CDS
; LOCATION: (29)..(1084)
US-09-282-305-3

Query Match          20.4%; Score 329.2; DB 4; Length 1475;
Best Local Similarity 63.8%; Pred. No. 1.8e-90;
Matches 499; Conservative 0; Mismatches 283; Indels 0; Gaps 0;

QY 528 caccatgcaagaagtcaggagcattgtcttctgttacgtcaatgatcgtctggcc 587
|||||
Db 164 caccacgccaagaagagcagcgctccggtctctgtctacgtcaacgacatcgtctcgcc 223
|||||

QY 588 atcctggaactgtctaaagtatcaccagaagggtgtgtacatgtgaattgatccaccat 647
|||||
Db 224 atcctgagctcctcaagttccacagcggtgtgtatatagtagaacattgtgtccaccat 283
|||||

QY 648 ggtgacggcgtggaagagcgcttcacacacagcagcggtcatgtactgttcttcat 707
|||
Db 284 ggagatggtgtgagagcgcttccttcaactacaacagagtcacgtactgttcttccac 343
|||||

QY 708 aagtatggagatcttccaggaactgtggagacctacgagataccggggtgcaagac 767
|||||
Db 344 aagtatggagatcttccaggaactgtggagacctacgagataccggggtgcaagag 403
|||||

QY 768 aagtatatgtgttaactaccgcgtccagagcgggattgtatgacgagtcctatgagcc 827
|||||
Db 404 aagcatatgtcttgaaatgttcccttgagtgatgtatcgtatgacacaccccttctgtgt 463
|||||

QY 828 atttcaagcgggtcatatgtccaaagttaatgagatgttccagcctagtgcgtgtctta 887
|||
Db 464 ctgttcaatgcatcatatgaagaagttatgaggtttatcagccagacgtgtgtctc 523
|||||

QY 888 cagtgtgctcagactccctatctggtgagtcgttaggtgtgtcttaactatctatcaaa 947
|||
Db 524 caatgcgagctgactccttgtgtgagagcaggttaggtgtgtcttaacacctgtctgtgaag 583
|||||

QY 948 ggaacgccaagtgtgtgaattgttcaagagccttcaacctgatactgtatgtgga 1007
|||
Db 584 ggtcatgtgactgcctccgttctccttagtgcgtacaatgttccatgatgtgttaggt 643
|||||

QY 1008 ggcggtgttacacacatctcgttaacgttgcgggtgtgtgacataatgagacagctgtggcc 1067
|||
Db 644 ggtgaggttacacacatcagaatgttgcacgctgtgtgtgtctacgagacgagctgtgt 703
|||||

QY 1068 ctgatacgagatccctaagtgttcacataatgacttaaccagaacacgaatgagtaactg 1127
|||
Db 704 gttgaggttgaacctgtataacaagctgtccttaacatgattactatgagtaacttggccct 763
|||||

QY 1128 gatttcaagctccacatcagctccttccaatatgacttaaccagaacacgaatgagtaactg 1187
|||||
Db 764 gattatacttctcatatccaacaaaagtgttgaaaaacctgaataccacaagaagactgtg 823
|||||

QY 1188 gagaagatcaaacagcgactgtttgagaaccttagaatgtgtccgacgcaactgtgggtc 1247
|||||
Db 824 gagaacataaagaacatgatatttgagaacctgttcaagaatagaacatgttccagcact 883
|||||

QY 1248 caaacgagcgatctcgtagagacgccaatccctgagagagtggtgagtgagagcgaagac 1307
|||
Db 884 caatccatgacagacgcgtcagacccttgaagctccagagagagaagagagagacatgtgac 943
|||||

QY 1308 ga 1309
```



```

; REFERENCE/DOCKET NUMBER: 52342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-770-379-20

```

Query Match	3.1%;	Score 50;	DB 2;	Length 32207;
Best Local Similarity	49.2%;	Pred. No. 0.00018;		
Matches 131; Conservative	0;	Mismatches 135;	Indels 0;	Gaps 0;

QY	1293	galtgagycgaagaacgacccctgacaaagcgcattctcgattctgtctctctgacaaacgaatt	1352
Db	20976	GACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGAT	20917
QY	1353	gcctgtgaggaagagtctctccgattcttgaagagggagagggggccgcaagaactct	1412
Db	20916	GACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGAT	20857
QY	1413	tccaacttcaaaaaagccaagagagtccaataacagagatgtaaaaaagaaagacccaag	1472
Db	20856	GACGAGGAGGATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	20797
QY	1473	gagaaagaaaggaatcacccgaagagagaaaaaaccaagagagagaaagccagaagccaaaagg	1532
Db	20796	GAGGACGAGGAGGATGACGATGATGAGGACCAATGAGGACGAGGAGGATGACGAGGAGGAG	20737
QY	1533	gtcaaggaaggaagccaagtggcctg	1558
Db	20736	GACCAAGAAGGAGGAGGAGGAGGAGCG	20711

RESULT 13
US-08-757-669A-20/c
; Sequence 20, Application US/08757669A
; Patent No. 6183751
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757, 669A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28, 678
; REFERENCE/DOCKET NUMBER: 45185-F
; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-757-669A-20

```

Query Match	3.1%;	Score 50;	DB 4;	Length 32207;
Best Local Similarity	49.28;	Pred. No. 0.00018;		
Matches 131; Conservative	0;	Mismatches 135;	Indels 0;	Gaps 0;

QY	1293	gacgaagcgaagaacgcaccctgacaagcgcatctcgatctgtctcccttgacaacgnaatt	1352
Db	20976	GACGAGGAGGATGACGAGGAGGATGACGAGCAGAGGATGACGAGGAGGATGACGAGGAGGAT	20917
QY	1353	gccctgtaggaagagtctctccgatcttgaaggaggaggaggggccgcaagaactct	1412
Db	20916	GACGAGGAGGATGACGAGGAGGATGACGAGCAGAGCATGACGAGGAGGATGACGAGGAGGAT	20857
QY	1413	tccaacttcaaaaaagccaagagagttcaaaacagaggtgaaaaagagaagaccacagag	1472
Db	20856	GACGAGGAGGATGACGAGGAGGAGGAGCAGCAGCAGCAGCAGCAGGAGGAGGAGCAGCAGCAGCAG	20797
QY	1473	gagaagaaaaggaatcacccgaagagagaaaaaaccaaggaggaagaagccagaagccaaaggg	1532
Db	20796	GAGGACGAGGAGGATGACGATGATGAGGACAATGAGGACGAGGAGGATGACGAGGAGGAG	20737
QY	1533	gtcaaggaaggagccaagtgtgacctg	1558
Db	20736	GACCAAGAAGGAGGACGAGGAGGACGG	20711

```

RESULT      14
US-09-230-371A-20/C
; Sequence 20, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 32207
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-20

```

Query Match	3.1%;	Score 50;	DB 4;	Length 32207;
Best Local Similarity	49.2%;	Pred. No. 0.00018;		
Matches 131; Conservative	0;	Mismatches 135;	Indels 0;	Gaps 0;
QY 1293	gatgagacgaagacgacccctgcacaagcgcatctcgatctgctcctctgacaaagaatt	1352		
Db 20976	GACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGAT	20917		
QY 1353	gcctgtgaggaagaattctctccgattctgaagaggaagggaagggggccgcaagaactct	1412		

Db	303	TTGGTGAGGACTGTCCAGTATTTCGATGCGCCTGTTTGGATTCTGTCAAGTTGTCTACTGCTG	362
QY	457	gttcctgtgcaagtgctgtgaaacttaataagcagcagacgacatcgccgtgaattgg	516
Db	363	GTTCTGTGGCAAGTGTCTGTCAACTTAATAAGCAGCAGACGACATCGCTGTGAATTGGG	422
QY	517	ctgggggacctgcaccatgcaagaagtcgcagagcatctggtctctgttacgtcaatgata	576
Db	423	CTGGGGGCCCTGCACCATGCAAGAAGTCCGAGGCATCTGGCTTCTGTACGTCATATGATA	482
QY	577	tcgtcttgccatctcctggaactgtctaaagtatcaccagaaggtgtctgtacattgacattg	636
Db	483	TCGTCTTGCCATCTCTGGACTGCTAAAGTATCACCCAGAGGCTGCTGTACATTGACATTG	542
AY	637	atatcaccatggtgagcgcgtggaagagccctctacaccaacygaccggtgcatgactg	696
Db	543	ATATTACCATGTGTGACGGCGTGGAGAAGGCCCTTCTACACCACGGACCGGGTCAATGACTG	602
QY	697	tgtcccttcataagatgtagaagtaactlcccaagaaactgggaacctacgagataccggg	756
Db	603	TGTCCTTTCATTAAGTATGAGAGAGTACTTCCACAGGAACCTGGGACCTACGGGATATCGGG	662
QY	757	ctggcaaaagacaagtatatgctgttaactaccgcctccgagacggaattgatgacgagt	816
Db	663	CTGGCAAAAGGCAAGTATTATGCTGTTAACTACCCGCTCCGAGACGGGATTGATGACGAGT	722
QY	817	cctatgagccatttccaagccggtcatgtlccaaagtaatgagatgtltccagccatgtg	876
Db	723	CCTATGAGGCCATTTTCAAGCCGGTCAATGTCCAAAGTAATGAGATGTTCACGCCCTAGTG	782
QY	877	cggtgtgtcttacagtggtggtcagactcccatctctgsgagatcgtttagtgttcttcaatc	936
Db	783	CGGTGGTCTTACAGTGTGGCTCAGACTCCCTATCTGGGGRTCGGTAGGTTGCTTCAATC	842
QY	937	tatctatcaaaagacacgcgccaagtgtgtt-ggaattgttcaagagctttaacctgacctatg	995
Db	843	TAACTATCAAAAGGACACGCCCAAGTGTGTGGGAATTGTCAAGAGCCTTTAACCTGCCTATG	902
QY	996	ctgattgctgggagcg 1012	
Db	903	CTGATGCTGGGAGGSGG 919	

RESULT	3	
AL546378		
LOCUS		
DEFINITION	AL546378	910 bp mRNA linear EST 16-FEB-2001
	AL546378 LTI_NFL006_PL2	Homo sapiens CDNA clone CS0DI031YC03 5
prime, mRNA sequence.		
ACCESSION	AL546378	
VERSION	AL546378.1	GI:12879433
WORDS	EST.	
URCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 910)	
AUTHORS	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.	
TITLE	Full-length CDNA libraries and normalization	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Genoscope	
	Genoscope - Centre National de Sequencage	
	BP 191 91006 EVRY cedex - France	
	Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.	
FEATURES	Location/Qualifiers	
source	1..910	

/note-"Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>

Query Match	54.4%;	Score 876;	DB 9;	Length 910;	
Best Local Similarity	99.0%;	Pred. No. 8.1e-172;			
Matches	901;	Conservative	1;	Mismatches	6;
				Indels	2;
				Gaps	2;

Db	841	TCCTCTGACAAACGAATTCCTGTGTAGGAGAGAGTTCTCCGATTTCTGAAGAGAGAGAGAG	900
QY	1395	ggggggccgca 1404	
Db	901	ggggggccgca 910	
RESULT	4		
LOCUS	AL541430	898 bp	mRNA
DEFINITION	AL541430 LTI_FL002_PL1	Homo sapiens	CDNA clone CS0DE006YL03 5 prime
ACCESSION	AL541430		
VERSION	AL541430.1	GI:12872494	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.		
TITLE	Full-length CDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Genoscope Genoscope - Centre BP 191 91006 EVRY cedex - France Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
FEATURES			
source	1.898	Location/Qualifiers	
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DE006YL03"		
	/clone_1lb="LTI_FL002_PL1"		
	/lab_host="DH10B"		
	/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand CDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded CDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"		
BASE COUNT	227 a	204 c	237 g 224 t 6 others
ORIGIN			
Query Match	52.6%	Score 847.6;	DB 9; Length 898;
Best Local Similarity	98.1%	Pred. No. 6.4e-166;	
Matches	883; Conservative	5; Mismatches	9; Indels 3; Gaps 3;
QY	104	gagcaagatgycgcagacgacgagcagccgaggaagtctgttactactacgagcgga	163
Db	1	GAGCAAGATGGCGCAGACGCGACCCGAGGAAGTCTGTACTACTACGACGGGGA	60
QY	164	tgttggaaattactattatgacaaagccacccaatgaagcctcaacgaatccgcgatgc	223
Db	61	TGTTGAAATTACTATTATTAAGACAAGGCCACCAATGAAGCCTCACCGAATCCGCATGAC	120
QY	224	tcataattgtctgtcctaactatgtgtctctacccaagaatgaatctatcgccctcaaca	283
Db	121	TCATAATTGTCTGTCACTATGTGTCTTACCGAAATGAATCTATCGCCCTCACAA	180
QY	284	agccaatgtctgagagatgaqccaagtlaccacagcgatgactacattaaattcttgctc	343
Db	181	AGCCAAATGCTGAGGAGATGACCAAGTACCACAGCGATGACTACATTAATTTCTGCGCTC	240
QY	344	catccgtccagataacacatgttggagtlacagcaagcgagatgcagagattcaacgttgta	403
Db	241	CATCCGTCAGATAACATGTTGGAGTACAGCAAGCAGATGCAAGATTCAACGTTGGTGA	300
QY	404	ggactgtccagtatctcgatgagcctgtttgagttctgtcagttgtctactgtgtttcgt	463

Db	Accession	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
Db	301	GGACGTGCCAGTATTCGATGCGCTGTTTGAGTTCGTGCTACTGGTGTCTGT	360											
QY	464	ggcaagtgctgtgaaactaataaagcagcagcagacatcgccgtgaattggtggtggg	523											
Db	361	GGCAAGTGCTGTGAACCTTAATAAGCAGCAGACGACATCGCTGTGAATTGGGCTGGGG	420											
QY	524	cctgcacccatgcaaaagaagtcgagggcatctggtctgttcaactgatatacgtcctt	583											
Db	421	CCTGCACCATGCAAAAGAAGTCGAGGCATCTGGCTTCCTGTTACGTCATGATATCTCTT	480											
QY	584	ggccatcctggaactgctaaaglatcacccagaggtgtgtacattgacattgatatcca	643											
Db	481	GGCCATCCTGGAACCTGTAAGTATCACACAGAGGGTGGTGTACATTGACATTTGATTCA	540											
QY	644	ccatggtgacggtgtggaagagggccttctacaccagacccgggtcatgactgtgtcctt	703											
Db	541	CCATGGTGACGGCGGTGGAAGAGGCCCTCTACACCACGAGCCGGGTGATGACTGTCTCTT	600											
QY	704	tcataagtatgagagatcttccccaaggaactggg-acctacggtataccggggtgtgca	762											
Db	601	TCATAAGTATGGAGAGTACTTCCCAAGAACTGGGNNACCTACGGGATATCGGGGCTGGCA	660											
QY	763	aagacaagtattatgctgtttaactaacccgctccgagaacgggattgatgacgagtcctatg	822											
Db	661	AAGGCAAGTATTATGCTGTTAACTACCCGCTCCGAGACGGGATTGATGACGAGTCTATG	720											
QY	823	aggccatttcaagccggtcatgtccaaagtaatgagatgttccagcctagtcggtg	882											
Db	721	AGGCCATTTTCAAGCCGGTCTWT-TTWAAGTAATGAGTTTTCACAGCCTAGTGGGTG	779											
QY	883	tcttacagtggtgctcagactccctatctggtgacggttaggttgcctcaatcatcta	942											
Db	780	TCTTACAGTGTGGCTCAGACTCCCTATCTGGGGATCGTTAGGTTGCTTCAATCTAACTA	839											
QY	943	tcaaagacacgccaagtgtgtgaattgtcaagagcttaacctgcctatgctgatgc	1002											
Db	840	TCAAAGGACACGCCAAGTGTGTGRRATTTGTCAAGAGCTTTAA-CTGCCATATGCTGATGC	898											
RESULT	5													
AL559849		AL559849	981 bp	mRNA	linear	EST 16-FEB-2001								
LOCUS														
DEFINITION		AL559849 LTI_FL011_BC1 Homo sapiens CDNA clone CS0DG005YG21 5 prime												
ACCESSION		AL559849												
VERSION		AL559849.1	GI:12905737											
KEYWORDS		EST.												
SOURCE		human.												
ORGANISM		Homo sapiens												
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;												
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.												
TITLE		1 (bases 1 to 981)												
JOURNAL		Li, W.B., Gruber, C., Jesse, J. and Polayes, D.												
COMMENT		Full-length cDNA libraries and normalization												
		Unpublished (2001)												
		Contact: Genoscope												
		Genoscope - Centre National de Sequencage												

Library was constructed by Life Technologies. Contact :
Feng Liang Life Technologies, a division of Invitrogen
9800 Medical Center Drive Rockville, Maryland 20850, USA
Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 271 a 226 c 267 g 213 t 4 others
ORIGIN

Query Match 52.6%; Score 847.4; DB 9; Length 981;
Best Local Similarity 97.5%; Pred. No. 7.2e-166;
Matches 900; Conservative 3; Mismatches 14; Indels 6; Gaps 4;

QY 643 accatggtgacgagcgtggaagagcctctacaccacgagccggtcatgactgtctcct 702
Db 1 ACCATGgtGACGCGCTGGAAGAGCCCTTACACACGACCGGCTCATGACTGTCTCT 60
QY 703 ttcataagatgagagtagtactcccaagaaactgggacctaagagataccgggctgca 762
Db 61 TTCATAAGTAATGGAGAGTACTTCCACGAACT-GGGACCTACGGGATATCGGGGCTGGCA 119
QY 763 aagacaagtattatgtctgttaactaccgcctccgagacgaggtatgatagcagagtcctatg 822
Db 120 AAGGCAAGTATATGCTGTTAACACCCTCCGAGACGGGATTGATGACGAGTCTCTATG 179
QY 823 aggccatttcaagccggtcatgttccaaagtaatgagatgttccagcctaagtgcgtgg 882
Db 180 AGGCCATTtTCAAGCCGGTCATGTCCAAGTAATGGAGATGTTCCAGCCTAGTGCCTGGTG 239
QY 883 tcttacagtggtgctcagactccctatctggsgatcggttagtggtcttcaatctatcta 942
Db 240 TCTTACAGTGTGCTCAGACTCCCTATCTGGGATCGGTTAGGTGCTTCAATCTAATACTA 299
QY 943 tcaagagacagcgaagtgtgtgaatttgtcaagagctttaacctgcctatgtctgtagc 1002
Db 300 TCAAGAGACACGCGCAAGTGTGTGAATTTGTCAAGAGACTTTAACCTGCCATATGCTGATGC 359
QY 1003 tgggagcggtggtttacaccatctgtaacgttgcgggtgtctgacatatgagacagctg 1062
Db 360 TGGAGCGCGGTGTTACACCATTCGTAACTGTGCCCCGGTCTGGACATATGAGACAGCTG 419
QY 1063 tggccctgtagatcgagatccctaataatgagcttccataacaatgactacttgaataacttg 1122
Db 420 TGGCCCTGATATCGAGATCCCTAATGAGCTTCCATATCAATGACTACTTTGAATACTTTG 479
QY 1123 gaccagatttcaagctccacatcagtccttccaatalatgactaaccagaaacgaatgagt 1182
Db 480 GACCAGATtTCAAGCTCCACATCAGTCTCTCAATATGACTAACCAGAACGAAATGAGT 539
QY 1183 acctggaagaatcaaacagcgactgtttgagaaccttagaatgtctgcgcgacgcacctg 1242
Db 540 ACCTGAGAAGATCAAAACAGCGACTGTTGAGAACTTAGAATGCTGCCGACGCACCTG 599
QY 1243 ggttccaaacgacgagcgattctctgagagcgcatccctgagagagtgcgagatgagagcg 1302
Db 600 GGGTCCAATATGACGCGCATTCCTGAGGACGCCATCCCTGAGGAGAGTGGCGATGAGGACG 659
QY 1303 aagacgacctgtaacaagcgcatctcgatctgtctcctctgacaaacgaatgtcctgtgag 1362
Db 660 AAGACGACCCCTGACAAAGCGCATCTCGATCTGCTCTCTGACAAACGAATTCCTGTGAGG 719
QY 1363 aagagtctccgattctgaagagagagggggcgcaagaactcttccaacttca 1422
Db 720 AAGAGTCTCTCGATTCCTGAAGAGGAGGAGGGGGCGCAAGAACTTCCAACTTCA 779
QY 1423 aaaaagccaagagagtcaaaacagagatgaaaaaagagaagaagaccagagagagaagaag 1482
Db 780 AAAAAAGCCAAGAGAGTCAAAACAGAGAGATGAAAAAGAGAAAGACCAGAGAGAGAAGAAG 839
QY 1483 gaatcaccgaagagagaaaaaaccaagagagagaagccagaagccaaggggttcaagagag 1542
Db 840 AAGTCACCCGAAGAGAG-AAAACCAAGAGAGAGAAC--AGAAGCCAAGGGGTCAAGAGAGA 896

QY 1543 agccaagtgtgcctgaatgagc 1565
Db 897 GGTC--ARKTGSCTGAATGAC 917

RESULT 6
AL558916 847 bp mRNA linear EST 16-FEB-2001
LOCUS
DEFINITION AL558916 LTI_NFL008_TC2 Homo sapiens cDNA clone CS0DJ007YP20 5
prime, mRNA sequence.
ACCESSION AL558916
VERSION AL558916.1 GI:12903904
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 847)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. 847
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DJ007YP20"
/clone_lib="LTI_NFL008_TC2"
/sex="male"
/tissue_type="T cells from T cell leukemia"
/note="Vector: pcwvSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-Oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pcwvSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 217 a 195 c 224 g 207 t 4 others
ORIGIN

Query Match 52.1%; Score 839; DB 9; Length 847;
Best Local Similarity 99.1%; Pred. No. 3.9e-164;
Matches 839; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 101 ggcgaagaatgycgcagacgagcgcacccggaggaagtctgttactactacgagc 160
Db 1 GCGGAGCAAGATGGGBCAGACGCGAGCGCACCCGGAGGAAAGTCTGTACTACGACGG 60
QY 161 ggaatgtgaaattactattatgacaagggcccaaatgaagcctcaacgaatccgcat 220
Db 61 GGATGTTGGAATTACTATTATGACAAAGGCCCAATGAAGCCTCACCGAATCCGCAT 120
QY 221 gactcataattgtctgctcaactatgtctctacccgaaaaatggaatctatcgccctca 280
Db 121 GACTCATATTTGCTGCTCAACTATGTGCTTACCGAAAAATGGAATCTATCGCCCTCA 180
QY 281 caaagccaatgtcgtgagagatgaccaagtaaccaagcgatgactacataaattctgcy 340
Db 181 CAAAGCCAATGCTGAGAGAGATGACCAAGTACCAACAGCGATGACTACATTAATCTTGGC 240
QY 341 ctccatccgtccagataacatgctcgaggtacagcaagcagatgcagagattcaacgttgg 400
Db 241 CTCATTCGCTCCAGATATACATGTCCGAGGTACAGCAAGCAGATGCAGAGATTCACGTTGG 300
QY 401 tgaggactgtccagtatctgatgagcgtgttgagttctgtcagttgtctactactgtgttc 460

Db 301 TGAGGACTGTCCAGTATTGCATGAGCCCTGTTGAGTTCGTCTGCTACTGTTGCTC 360
QY 461 tgtgcaagtgtgtgaacttaataagcagcagacatcgccgtgaattggtg 520
|||||
Db 361 TGTGGCAAGTGTGTGAACCTAATAAGCAGCAGACGACATCGCTGTGAATGGGCTGG 420
QY 521 gggccctgcaccatgcaagaagtcgcagagcatctgtctgttaacgtcaatgatatacgt 580
|||||
Db 421 GGGCCCTGCACCATGCAAGAAGTCCGAGGCATCTGGCTTCTGTACGTCAATGATATCGT 480
QY 581 ctggccatccctggaactgttaaatatcacccagaggtgtgtacattgacattgatat 640
|||||
Db 481 CTTGGCCATCCTGGAACCTGAAGTATCACCAGAGGGGTGTACATGACATGATATAT 540
QY 641 tcaccatgtgacggtgtgaaagagccttctacacacagcagcgggtcatgtgtc 700
|||||
Db 541 TCACCATGTGACGGCGGTGAAGAGGCGCTTCTACACACGACCGGGTCACTGTGTCTC 600
QY 701 cttcataagtatgagagtaacttcccaaggaactggggaacctacgggataccgggctgg 760
|||||
Db 601 CTTTCATAGTATGAGAGTACTTCCAGGAACCTGGGAGCTACGGGATATCGGGGCTGG 660
QY 761 caaagacaagtattatgtgttaactaacccgtccagagacgggattgtagcagagtccta 820
|||||
Db 661 CAAAGGCAAGTATTATGCTGTAACTACCCSCTCCSAGACGGGATGATGACGAGTCTTA 720
QY 821 tgaagccatttcaagccgtcatgtccaaagtaatgagatgttccagcctagtcggt 880
|||||
Db 721 TGAGGCCATTTTCAAGCGGTCTATGTCCAAAGTAATGAGATGTTCCASCTAGTGGGT 780
QY 881 ggtctacagtgtgtcagactccctatctggtggtggttagttgcttcaatcattc 940
|||||
Db 781 GGTCTTACAGTGTGCTCAGACTCCCTATCTGGGGTGGGTGAGTTGCTTCAATCTAAC 840
QY 941 tatcaaa 947
|||||
Db 841 TATCAAA 847

RESULT 7
AL522933 832 bp mRNA linear EST 13-FEB-2001
LOCUS AL522933 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB009YN22 5
DEFINITION prime, mRNA sequence.
ACCESSION AL522933
VERSION AL522933.1 GI:12786426
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 832) Jesse, J. and Polayes, D.
AUTHORS Li, W.B., Gruber, C.,
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Source location/Qualifiers
1. 832
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB009YN22"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed

by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com>
BASE COUNT 236 a 194 c 227 g 175 t
ORIGIN

Query Match 51.1%; Score 824; DB 9; Length 832;
Best Local Similarity 99.4%; Pred. No. 5e-161;
Matches 827; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 773 ttatgctgttaactaccgctccgagacggattgatacyagtccatagggcatttt 832
|||||
Db 1 TTATGCTGTTAACTACCCGCTCCGAGACGGGATGATGACGAGTCCATGAGGCCATTTT 60
QY 833 caagccggtcatgtccaaagtaatgagatgttccagcctagtgcgtgtcttacagt 892
|||||
Db 61 CAAGCCGTCATGTCCAAAGTAATGAGATGTTCCAGCCTAGTGGGTGCTTACAGTG 120
QY 893 tggctcagactccctatctggggtatcggttagttgtctcaatctatctatacaaggaca 952
|||||
Db 121 TGGCTCAGACTCCCTATCTGCGGATCGGTAGGTGCTTCAATCACTAATCAAGAGACA 180
QY 953 cgccaagtgtgtgaattgtcaagagctttaacctgtctatgtctgtggaagcgg 1012
|||||
Db 181 CCCCAGTGTGTGGAATTTGTCAAGAGCTTTAACCCTATGCTGATGCTGGGAGCGG 240
QY 1013 tggttacaccattcgttaacgtttgccggtgtgtgacatatgagacagctgtgcccctgga 1072
|||||
Db 241 TGGTTACACCATTCGTAACTGTGCCCGGTGCTGACATATGAGACAGCTGTGGCCCTGGA 300
QY 1073 taaggagatccctaataatgacttccatacaaatgactacttgaatacttbgaccagattt 1132
|||||
Db 301 TACGGAGATCCCTAATGACTCTCCATACAAATGACTTACTTTGAATACTTTGACCAGATTT 360
QY 1133 caagctccacatcagtccttccaaatatagtactaaccaagaacaaatgacttgagaa 1192
|||||
Db 361 CAAGCTCCACATCAGTCTCTCCAAATATGACTAACCAAGAACGAATGAGTACTGGAGAA 420
QY 1193 gatcaaacagcagactgtttgagaaccttagaatgtgtccgcaacgacactgggtccaac 1252
|||||
Db 421 GATCAAAACAGCGACTGTTTGAGAAACCTTAGAATGCTGCCGACGACACTGGGGTCCAAT 480
QY 1253 gcaagcgattcctgtgagagcgcacccctgagagagtggtgagagcagaacgcc 1312
|||||
Db 481 GCAGGCGATTCTCTGAGAGCGCCCATCCCTGAGGAGAGTGGCGATGAGGACGAAGCAGACC 540
QY 1313 tgacaagcgcacatcgcattgtctctctgtgacaacagaattgctgtgaggaagattctc 1372
|||||
Db 541 TGACCAAGCGCATCTCGATCTGCTCTCTGACAAACGAATTGCCCTGTGAGGAAGATTCTC 600
QY 1373 cgattctgaagagagagagggggccgcaagaactcttccaactcaaaaaagccaa 1432
|||||
Db 601 CGATTCTGAAGAGAGAGGGGGCGCAAGACTCTTCCAACTTCAAAAAAGCCAA 660
QY 1433 gagaagcaaacagagatgaaaaagaaagaaagcagagagagaaagaaagaaatcacgga 1492
|||||
Db 661 GAGAGTCMAAAACAGAGAGATGAAAAAGAAAAGAACCCAGAGAGAAAGAAAGATCACCGA 720
QY 1493 agagagaaaaccaagagagaaagccagaagccaaagggttcaagagagggccaaglt 1552
|||||
Db 721 AGAGGAGAAAACCAAGGAGAGAGAGCCAGAGCCCAAGGGGTCAAGGAGAGGTCAAGTT 780
QY 1553 ggcctgaatgagacctctccagctctgtgcttctgtgagtcctcctacggttcc 1604
|||||
Db 781 GGCTGAATGAGACTTCTCCAGCTCTGTGCTTCTGTGAGTCCCTCACGTTTC 832

RESULT 8
BC013141 1627 bp mRNA linear HTC 29-AUG-2001
LOCUS BC013141

RESULT 9
LOCUS BG754715 954 bp mRNA linear EST 15-MAY-2001
DEFINITION 602714475F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4854547 5',
mRNA sequence.
ACCESSION BG754715
VERSION BG754715.1 GI:14065368
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 954)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1702 row: f column: 20
High quality sequence stop: 915.
Location/Qualifiers
1. 954

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4854547"
/clone_lib="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC library."

BASE COUNT 249 a 214 c 265 g 226 t
ORIGIN

Query Match 49.0%; Score 789.2; DB 10; Length 954;
Best Local Similarity 94.0%; Pred. No. 8.8e-154;
Matches 897; Conservative 0; Mismatches 48; Indels 9; Gaps 7;

QY 96 ggggagcgagcaagatggcgacagcgagcgaccggaggaaagtctgttactactac 155
|||||
Db 1 GGGGAGGCGAGCAGATGGCGCAGACGCA-GGCACCCGAGGAAGTCTGTACTACTAC 59
QY 156 gacgggagatgttgaatactattatggacaaggcccaaatgaagcctcaccgaatc 215
|||||
Db 60 GACGGGATGTTGGAATTAATTAATATGATGACCAAGGCCCAATGAAGCCTACCGAATC 119
QY 216 cgcattgactcataattgtctctcaa-ctatgtctctaccgaaanaatggaatctatcg 274
|||||
Db 120 CGCATGACTCATATTCTCTGCTCAAGCTATGGTCTCTACCGAAAAATGGAATCTATCG 179
QY 275 ccctcacaagccaatgtctgagagatgacccaagtcacacagcgatgactacataaatt 334
|||||
Db 180 CCCTCACAAGCCAATGCTGAGGAGATGACCAAGTACACACGCGATGACTACATAAATT 239
QY 335 ctgagctccatccgtccagataacatgtctcgagtaacagcaagcagatgcagagattcaa 394
|||||
Db 240 CTTGCGCTCCATCCGTCAGATAACATGTCGAGATACAGCAAGCAGATGCAGAGATTCAA 299
QY 395 cgttgtgagagactgtccagtatcgtatgagcgctgttgaattctgtcagttgttactactg 454
"

Db 300 CGTTGGTGAGGACTGTCCAGATATTGCATGGCCCTGTTGAGTTCTGTCAAGTTGTACTCGG 359
QY 455 tggctctgtgcaagtgtctgtgaacttaataagcagcagcgacatcgccgtgaattg 514
|||||
Db 360 TGGTTCGTGTGGCAGAGTGTGTGAACCTTAATAAGCAGCAGACGACATCGCTGTGAATTG 419
QY 515 ggcctggggcctgcaccatgcacaagaagtcgcgagcgatctgtctgttaacgtcaatga 574
|||||
Db 420 GGCTGGGGCCCTGCACCATGTCAAGAAGTCCGAGGCATCTGGCTTCTGTACGTCAATGA 479
QY 575 taccgtcttggccatcccttgaactgtctaaagatcaccaagagggtgtgtacattgacat 634
|||||
Db 480 TATCGTCTTGCCATCTGGAAGTCTTAAGTATCACACAGAGGGTGTGTACATTGACAT 539
QY 635 tgataatcaccatgtgtgacggtgtggaagaggccttctacacacagcagcggtcatgac 694
|||||
Db 540 TGATATTACCATGATGATGACGGCGTGGAAGAGGCGCTTGTACACACGAGCGGTCATGAC 599
QY 695 tgtgtccttcataaagtatgagagtaacttccaccaggaactggggacctaagcagccgg 754
|||||
Db 600 TGTGTCTTTCATTAAGTATGAGAGTACTTCCACAGAACTGGGGACCTACGGGATGATCGG 659
QY 755 ggcctggcaagaacaagtattatgtctgttaactaccgcctccagagacggatga-tgacg 813
|||||
Db 660 GGCTGGCAAAAGCAAGTATATGCTGTTAACACTACCCGCTCCGAGACGGGATGATGACG 719
QY 814 agtccatagagccatttccaagccggtcatgttccaa--gtaatgagatgttccag-- 869
|||||
Db 720 AGTCCATGAGGCCATTTTCAAGCCCGTCAATGTCCAAAGTAATGAGATGTTCACAGC 779
QY 870 cctaagtcggtgtgtcttaacgtgtg-gtcacagctccctatcttgggacgtgtagttg 928
|||||
Db 780 CTAGATGCGGTGTCTAACAGGGTGAGCTCAGAACTCTATCTGGGGATCCGCAAGGTGG 839
QY 929 ctccaatctatcatcaagaagacacgccaagtgtgtgaatttgcagaagcttaacct 988
|||||
Db 840 TTCACTCTCTAATATCAAGGAACGCCCAAGGTGTGTGCAATTGTCAAGAGC-TTAACTT 898
QY 989 gactatgctgatgtctggagcggtgtgttacaccatcgtgaacgttgcccggtg 1042
|||||
Db 899 GCCTATGCTGATGCTGAAGCGGCTGTAAACCATTTCTGTAACGTTGACCGTG 952

RESULT 10
LOCUS BG756668 910 bp mRNA linear EST 15-MAY-2001
DEFINITION 602715527F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4855663 5',
mRNA sequence.
ACCESSION BG756668
VERSION BG756668.1 GI:14067321
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 910)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1705 row: e column: 08
High quality sequence stop: 818.
Location/Qualifiers
1. 910

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4855663"
/clone_lib="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 231 a 201 c 253 g 224 t 1 others

ORIGIN

Query Match 48.4%; Score 779.4; DB 10; Length 910;
Best Local Similarity 95.7%; Pred. No. 9.4e-152;
Matches 833; Conservative 0; Mismatches 32; Indels 5; Gaps 3;

439 gtcaagtgtctactgtgtgtgtctgtgtgcaagtgtgtgaacttaataagcagcagc 498
|||||
2 GTCAGTTGCTACTGCTGCTTCTGTGGCAAGTCTGTGAACCTTAATAAGCAGCAGACGG 61
499 acatcgccgtgaattggcgtggcgccctgcacatgcaagaagtccgagcgatctgct 558
|||||
62 ACATCGCTGTGAATTGGCGCTGGGGCGCTGCACCATGCCAAGAAGTCCGAGGCATCTGGCT 121
559 tctgttacgtcaatgatatacgtctgtgccaatcctgtgaactgttaaaagtatcacccagag 618
|||||
122 TCTGTACGTCAATGATATCGTCTTGCCATCCTGGAAGTGTAAAGTATCACCAAGAGG 181
619 tgcgtacatgtacatgtatataaccatgtgtgacgagcgctgtgaagagcgcttctacacca 678
|||||
182 TGCTGTACATTTGACATTTGATATTCACCATGGTGACGGCGTGGAAGAGGCGCTTCTACACCA 241
679 cggaccggtgcatgactgtgtccttcataagatgagagtlactcccaggaactgggg 738
|||||
242 CGGACCGGGTCACTGTGTCTTTCATTAAGTATGGAAGTACTTCCAGGAACCTGGGG 301
739 acctaaggataccggggcgtgcaagaacataltatgtgttaactaccgcctccag 798
|||||
302 ACCTACGGGATATCGGGCGCTGGCAAGGCAAGTATATGCTGTTAACTACCGCGCTCCGAG 361
799 acgggattgtatgacgagtcctatgagccatttcaagccggtcatgtccaagaatag 858
|||||
362 ACGGATGTGATGACGAGTCCCTATGAGGCCATTTCAAGCCGTCATGTCCAAGTAATGG 421
859 agatgttccagcctagtgcgtgtgtcttacagtgtgtgtcagactccctatctgggac 918
|||||
422 AGATGTTCACGCTACTGCGGTGTGCTTACAGTGTGGCTCAGACTCCCTATCTGGGGATC 481
919 ggtaggtgtctcaatctatctatcaagaagacagcccaagtgtgtgaatttgcaga 978
|||||
482 GGTAGGTTGCTTCAATCTAATATCAAAAGACACGCCAAGTGTGTGAATTTGTCAAGA 541
979 gctttaacctgctatgtatgtgtgagagcggtgttacaaccattcgtaaagttggcc 1038
|||||
542 GCTTTAACCTGCTTATGTGATGTGGGAGGCGGTGTTACACCATTCGTTAAGCTTGCCC 601
1039 ggtgtgacatatgagacagcgtgtgcccgtgatacggagatccctaagcttccat 1098
|||||
602 GGTGCTGACATATGAGACAGCTGTGGCCCTGTGATACGGAGATCCCTAATGAGCTTCCAT 661
1099 acaatgactacttgaataactttgaccagagattcaagctccacatcagtccttccaata 1158
|||||
662 ACAATGACTACTTTGAATACTTTGAGCCAGATTTCAGCTCCACATCAGTCTTCCATTA 721
1159 tgactaaccaagaca--cgaatgagtacctgtgagagaagatcaaacagcgactgtttgaga 1216
|||||

Db 722 TGACTAACGAGACACGAATGAGTACCTGGAGAGATCAACAG-GACTGTTTGAGAA 780
QY 1217 ccttagaatgtccgcacgcacactgggtgtccaaacgcagg--cgatcctgagcagcc 1274
|||||
Db 781 CCTTAGAAATGCTTGGCGACGCGACTGGGGTCCAAATGCAGGGCGATTCCNTGAAGACGCC 840
QY 1275 atccctgagagagtgcgatgagcagaa 1304
Db 841 ATTCTGAGAGAGAGTGGGATGAGGACAA 870

RESULT 11
BM477303
LOCUS BM477303 982 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6485091 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5554070
5', mRNA sequence.
ACCESSION BM477303 GI:18526345
VERSION BM477303
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 982)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12272 row: a column: 15
High quality sequence stop: 677.

FEATURES Location/Qualifiers

source

1. 982
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5554070"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 251 a 225 c 269 g 235 t 2 others

ORIGIN

Query Match 48.3%; Score 778; DB 10; Length 982;
Best Local Similarity 96.7%; Pred. No. 1.9e-151;
Matches 847; Conservative 0; Mismatches 22; Indels 7; Gaps 5;
QY 97 gggagcgagcaagatggtcgagacgagcgagcccgaggaagtctgttactactag 156
|||||
Db 19 GGGAGCGAGCAAGATGGCGCAGACGCGAGGCCGCCGGAAGTGTGTTACTACTACG 78
QY 157 acgggagatgttgaataactatataygacaagggccaccaatgaagcctcaccgaatcc 216
|||||
Db 79 ACGGGATGTGGAATATCTATATGACAAAGCCACCAATGAAGCCCAACCGAATCC 138
QY 217 gcatgactcataattgtctgtcctaactatgtctctacccgaaanaatgtgaatctatgcc 276
|||||
Db 139 GCATGACTCATATTTGCTGCTCAACTATGTGTCTTACCAGAAAATGGAATCTATCGCC 198
QY 277 ctcaacaagccaatgctgaggaagatgaccaagtaaccaagcgatgactacattaattct 336
|||||

Db 199 CTCACAAAGCCATGCTGAGAGATGACCAAGTACCACAGCGATGACTACATTAAATCT 258
QY 337 tgcgtccatccgtccagataacatgtcggagtagcagcaagcagatgcagagattcaacg 396
Db 259 TCCGCTCCATCCGTCACAGTAACATGTCCGAGTAGACAGCAAGATGCAGAGATTCAACG 318
QY 397 ttggttaggactgtccagatctcgatggtcgtgttgagttctgtlcaagttgtctactgtg 456
Db 319 TTGGTGAGGACTGTCCAGTAATCGATGGCCGTGTGAGTCTGTCAAGTTGTCTACTGTG 378
QY 457 gtctgtgcaagtgtctgaaactaataagcagcagcagacatgcgcgtgaattggg 516
Db 379 GTTCTGTGGCAAGTGTGTGAACTTAATAAGCAGCAGACGACATCGCTGTGAATTGGG 438
QY 517 ctgggggctgcaccatgcgaagaagtcaggagcatctgtctgttacctgaatgata 576
Db 439 CTGGGGGCTGCACCATGCAAGAAGTCCGAGGCATCTGGCTTCTGTACGTCAATGATA 498
QY 577 tcgtcttgccatccctggaactgtctaagatacaccagaagggtgtgtacattgacattg 636
Db 499 TCGTCTTGCCATCCTGGACTGCTAAAGTATCACACAGAGGGTGTGTACATTGACATTG 558
QY 637 atattcaccatggtgacggtggaagagcgctctacaccacgagcggtcatcactg 696
Db 559 ATATTCAACCATGTTGACGGCGGTGGAAGAGCGCTTCTACACACGACCGGGTCATGATG 618
QY 697 tgcctctcataagtagtagagtagtctcccaagaaatgggagacctagcgagatccgggg 756
Db 619 TGTCTTTTCATAGTAGAGTAGACTTCCAGGAACGCGGACCTACGAGATATCGGG 678
QY 757 ctggcaagaacaagtattatgctgttaactaccgcgtccgagacgggattgatgagagt 816
Db 679 CTGGCAAAAGCAAGTATATCTGTTAACTACCCGCTCCGAGACCGGATTGATGACGAGT 738
QY 817 cctatgagcgccatttcaagccggtcatgttccaaa-gtaatgagatgtt-ccagcctag 874
Db 739 CCTATGAGGCCATTTCAGCCGTCATGTCCAAAGTATGAGATGTTCACAGCCCTAG 798
QY 875 tgcggt-ggtcttaccagt-gtggctcagactccctactctgggg---atcggttaggttgc 929
Db 799 TCGCGTGGGTCTTACAGTGTGGCTCAGACTCCCTATCTGNGGATCCGGTTAGGGTTGC 858
QY 930 ttcaatctatctatcaaaaggaacagccaagtgtgtg 965
Db 859 TTCAATCTAATCTATCAAGGAACACCCCAAGNTGTG 894

RESULT 12
BI772704 778 bp mRNA linear EST 25-SEP-2001
LOCUS 603053373F1 NIH_MGC_122 Homo sapiens CDNA clone IMAGE:5202739 5',
DEFINITION mRNA sequence.
ACCESSION BI772704
VERSION BI772704.1 GI:15764282
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 778)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM1508 row: b column: 20

High quality sequence start: 6
High quality sequence stop: 762.
Location/Qualifiers
1. 778
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5202739"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."

BASE COUNT 201 a 177 c 207 g 193 t
ORIGIN

Query Match 45.8%; Score 737.4; DB 10; Length 778;
Best Local Similarity 99.0%; Pred. No. 4.8e-143;
Matches 763; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 136 ggaagctctgttactactacgacgggagtggtggaattactattatgacaaagccacc 195
Db 8 GGAAGTCTGTACTACTACGACGGGATGTTGGAATTACTATTATGACAAAGGCCACC 67
QY 196 caatgaagcctcaccgaatcc-gcatgactcataattgtctcactatgctctac 254
Db 68 CAATGAAGCCTCACCGAATCCGGCATGACTCATTAATTGCTCTCAACTAATGCTCTAC 127
QY 255 cgaaaaatggaatctatgcctcacaagaagcaatgtagagatgacaaagtcaccac 314
Db 128 CGAAAAATGGAATCTATGCCCCCTCACAAAGCTAATGCTGAGAGATGACCAAGTACAC 187
QY 315 agcgatgactacattaattctgtgcgtccatccgtccagataacatgtcggagtagc 374
Db 188 AGCGATGACTACATTAATCTTGCGCTCCATCCGTCCAGATAACATGTCGAGTACAGC 247
QY 375 aagcagatgcagagattcaacgttgtgagagactgtccagtatctcgatgacctgttag 434
Db 248 AAGCAGATGCAGAGATTCAACGTGTGTGAGGACTGTCCAGTATTCGATGGCCGTGTGAG 307
QY 435 ttctgtcagttgtctactgtgttctgtgtgcaagtgtgtgaaacttaataagcagcag 494
Db 308 TTCTGTCAAGTTGTCTACTGTGGTGTCTGTGCGCAAGTCTGTGAACCTTAATAAGCAGCAG 367
QY 495 aaggaacatgcgcgtgaattggctggggcctgcaccatgcaaaagaagtcgaagcattct 554
Db 368 ACGGACATCGCTGTGAATGGGCTGGGGCCTGCACCATGCAAAAGAATCCGAGCATCT 427
QY 555 ggctctgttacgtcaatgatatcgcttggccatcctggaactgtctaagatcaccag 614
Db 428 GGCTTCTGTACGTCAATGATATCGTCTTGCGCATCTGGAACCTGTAAGTATACACAG 487
QY 615 aggtgtgttacattgacattgatatccaccatggtgacggcggtggaagagggccttctac 674
Db 488 AGGGTGTGTACATGACATTGATATTACCATGTGTGACGGCGTGAAGAGGCTTCTAC 547
QY 675 accacggaccgggtcatatgactgttctcttcataagtagtagagtagtactcccaagaa 734
Db 548 ACCACGGACCGGGTCAATGACTGTCTCTTTCATAGTATGAGAGTACTTCCAGGAAGT 607
QY 735 ggggacctacgggataccggggcgtggcaagaagatatattgtgttaactaccgcctc 794
Db 608 GGGGACCTACGGGATATCGGGGCTGGCAAAAGGCAAGTATTATGCTGTAACTACCCGCTC 667
QY 795 cgagaaggattgatgacagagtcctatgagggcatttcaagccggtcatgttcaaaagta 854
|||||

Db 668 CGAGACGGGATGTGATGACGAGTCTCTATGAGGCCA-TTTCAGCCGGTTCATGTCCAAGTA 726

QY 855 atggagatgttccagcctagtgcgtgtcttaccagtgtgtcagactcc 905
|||||
Db 727 ATGGAGATGTTCACAGCCTAGTCCGGGGTCTTACAGTGTGGCTCAGACTCC 777

RESULT 13
BM452740 1007 bp mRNA linear EST 05-FEB-2002
LOCUS BM452740
DEFINITION AGENCOURT_6401305 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5498480
5', mRNA sequence.
ACCESSION BM452740
VERSION BM452740.1 GI:18501780
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1007)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lou Staudt

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM12130 row: e column: 09
High quality sequence stop: 558.

FEATURES
Source
1.1007
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5498480"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 256 a 259 c 257 g 235 t
ORIGIN

Query Match 45.7%; Score 736.8; DB 10; Length 1007;
Best Local Similarity 97.5%; Pred. No. 6.8e-143;
Matches 770; Conservative 0; Mismatches 17; Indels 3; Gaps 2;

QY 405 gactgtccagttatcgtatggcctgtttgagttctgtcagttgttactgtgtgtctgtg 464
|||||
Db 1 GACTGTCCAGTATTCGATGGCCGTGTTGAGTTCTGTCAAGTGTCTACTGTGTTCTGTG 60
QY 465 gcaagtgtgtgaaacttaataacagcagcagacgacatcgccgtgaattggtggggc 524
|||||
Db 61 GCAAGTGTGTGAACCTTAATAACGACGACGACATCGCTGTGAATTGGGTGGGGC 120
QY 525 ctgcaccatgcaagaagtcgagagcatctgtctgtttagtcaatgatatcgtcttg 584
|||||
Db 121 CTGCACCATGCAAGAAGTCCGAGGCATCTGGCTCCTGTACGTCAATGATATCGTCTTG 180
QY 585 gccatctggaactgtctaaagtatcaccagaggggtgtgtacattgacattgatattcac 644
|||||
Db 181 GCCATCCTGGAAGTGTAAAGTATCACAGAGGGGTGTGTACATTGACATTGATATTAC 240
QY 645 catgtgtacggcggtggaagagccttctacaccagcagcgggtcatgtactgtgtcctt 704
|||||

Db 241 CATGTTGACGGCGGTGGAAGAGGCCCTTCTACACCACGACCGCGGTTCATGACTGTGTCCTTT 300

QY 705 cataagtatggagagttacttcccagaaacttgggacctacgggataccgggctggcaaa 764
|||||
Db 301 CATAAGTATGAGAGTACTTCCAGAGAACTGGGGACCTACCGGGATATCGGGGCTGGCAAA 360

QY 765 gacaagtattatgctgttaactaccgcctccgagacggagatgtatgacagagtcctatgag 824
|||||
Db 361 GGCAAGTATTATGCTGTTAACCTACCCGCTCCGAGACGGGATGTGACGAGTCCCTATGAG 420

QY 825 gccatttccaagccggtcatgttccaaagtaatgagatgttccagcctagtgcgtgtgc 884
|||||
Db 421 GCCATTTCGAAGCCGGTCATGTCCAAAGTAATGAGATGTTCCAGCCTAGTGGGTGGTC 480

QY 885 ttacagtgtgtcagactccctatctctgggatcggttagtgttcaatctatctatc 944
|||||
Db 481 TTACAGTGTGGCTCAGACTCCCTATCTGGGGATCGGTTAGGTGCTCAATCTAATATC 540

QY 945 aaagacacgccaagtgtgtggaattgtcaagagctttaaactgtcctatgtgtgctg 1004
|||||
Db 541 AAAGACACGCCCAAGTGTGTGAATTGTCAAGAGCTTTAACCTGCTATGCTGATGCTG 600

QY 1005 ggaagcggtgttaccacatctgttaacgttgcgggtgtgtgacatatgagacagctgtg 1064
|||||
Db 601 GGAAGCGGTGTTCACACCATTCGTACGTTGCCCGGTGCTGACATATGACACAGCTGTG 660

QY 1065 gccctgtacagcagagatccctaattgagcttcacatacaatgactacttgaacttga 1124
|||||
Db 661 GCCCTGATACGAGATCCCTAATGAGCCTCCATCAATGACTACTTGAATACTTTGGA 720

QY 1125 ccagattccaagctccacatcag-tccttccaatatgactaaccag--aacacgaatgag 1181
|||||
Db 721 CCAGATTTCAGCTCCACATCAGACCCCTCCCAATATGACTACCCGGAACCCCAATGAG 780

QY 1182 tacctggaga 1191
|||||
Db 781 TACCCCTGGGA 790

RESULT 14
BG768807 803 bp mRNA linear EST 15-MAY-2001
LOCUS BG768807
DEFINITION 602743069F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4872689 5',
mRNA sequence.
ACCESSION BG768807
VERSION BG768807.1 GI:14079460
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 803)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DFP

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM1749 row: j column: 18
High quality sequence stop: 801.

FEATURES
Source
1.803
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4872689"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"

/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library. |"

BASE COUNT	207 a	188 c	229 g	179 t
ORIGIN				

Query Match	45.58%	Score 732.4;	DB 10;	Length 803;
Best Local Similarity	99.18%	Pred. No. 5.3e-142;		
Matches 747; Conservative	0;	Mismatches 6;	Indels 1;	Gaps 1;

Y	97	gggaggcgagcaagatggcgccagagcgcagggcaccocggaggaaagtctgttactactacg	156
Db	50	gggaggcgagcaagatggcgccagagcgcagggcaccocggaggaaagtctgttactactacg	109
OY	157	acggggatgttggaaattactattatgacaagggccaccaatgaagcctcaaccgaatcc	216
Db	110	acggggatgttggaaattactattatgacaagggccaccaatgaagcctcaaccgaatcc	169
OY	217	gcatactcataaatttgcgtctcaactatgctctacccgaaaaatggaatctatcgcc	276
Db	170	gcatactcataaatttgcgtctcaactatgctctacccgaaaaatggaatctatcgcc	229
OY	277	ctcacaagccaatgctgagggatgaccaagtlaccacagcgatgactacatlaattct	336
Db	230	ctcacaagccaatgctgagggatgaccaagtlaccacagcgatgactacatlaattct	289
OY	337	tgcgctccatccgctccagataacatgctcgagtlacagcaagcagatgcagaagattcaag	396
Db	290	tgcgctccatccgctccagataacatgctcgagtlacagcaagcagatgcagaagattcaag	349
OY	397	ttggtgaggactgtccagtalccagatgagcctgttgaagttctgtcagttgtactgtgt	456
Db	350	ttggtgaggactgtccagtalccagatgagcctgttgaagttctgtcagttgtactgtgt	409
OY	457	gtctctggcgaagtgtctggaacttaaaagcagcagacgacatcgccgtgaattgg	516
Db	410	gtctctggcgaagtgtctggaacttaaaagcagcagacgacatcgccgtgaattgg	469
OY	517	ctgggggcctgcacccatgcaaaagatccagagcactggtctgttaagtcgaatgata	576
Db	470	ctgggggcctgcacccatgcaaaagatccagagcactggtctgttaagtcgaatgata	529
OY	577	tcgtcttggccatccctggaactcgtctaaagtatcacccagaggtgtctgtatcatgtacatg	636
Db	530	tcgtcttggccatccctggaactcgtctaaagtatcacccagaggtgtctgtatcatgtacatg	589
OY	637	atatccaccatggtgacgcgctgtygaagagggccttctacaccaacgagccgggtcatgactg	696
Db	590	atatccaccatggtgacgcgctgtygaagagggccttctacaccaacgagccgggtcatgactg	649
OY	697	tgctccttcataagatgagagagtaactccacgaactgaggacctaagcgataaccggg	756
Db	650	tgctccttcataagatgagagagtaactccacgaactgaggacctaagcgataaccggg	709
OY	757	ctggcaagacaagtattatgctgttaactacccgctccgagacggatgtgatgacgagt	816
Db	710	ctggcaagacaagtattatgctgttaactacccgctccgaaacgggatgtgatgacgagt	769
OY	817	ccatagagg-ccattttccaagccgggtcatgtcca	849
Db	770	cctatgagggcccatTTTCAAGCCGGTCATGTCCA	803

RESULT 15
BM472180

LOCUS	BM472180	1034 bp	mRNA	linear	EST 05-FEB-2002
DEFINITION	AGENCOURT_6469876	NIH_MGC_92	Homo sapiens	CDNA clone	IMAGE:5576396
	5', mRNA sequence.				
ACCESSION	BM472180				
VERSION	BM472180.1	GI:18521222			
KEYWORDS	EST.				
SOURCE	human.				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 1034)	NIH-MGC http://mgc.nci.nih.gov/ .	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.			

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM12328 row: c column: 21
High quality sequence stop: 590.

FEATURES	Location/Qualifiers
source	1. .1034

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5576396"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

BASE COUNT	267 a	262 c	283 g	222 t
ORIGIN				

Query Match	45.2%;	Score 728.6;	DB 10;	Length 1034;
Best Local Similarity	97.3%;	Pred. No. 3.4e-141;		
Matches 752; Conservative	0;	Mismatches 19;	Indels 2;	Gaps 1;

OY	97	gaggaagcgagcaagaatggcgcagacgcagggccaccgcgaaggaaagtctgttactactaacg	156
Db	21	GGGAGGCAGCAAGATGGCGCAGACGCAGGGGCCACC GGGAAGTCTGTACTACTACG	80
OY	157	acgggatgttgaaaattactatatatgacaagccaccaatgaagcctcacccaatcc	216
Db	81	ACGGGATGTTGAAATTACTATTATGACAAGCCACCCAATGAAGCCTCACCGAATCC	140
OY	217	gcatactcataattttgtctgtctaactatggtctctaccgaaaaatgyaaatcatcgcc	276
Db	141	GCATGACTCATTAATTGCTGCTCAACTATGGTCTCTAACGAAAAATGGAATCTATCGCC	200
OY	277	ctcacaaagccaatgctgagagatgacccaagtlaccacagcgatgactacattaattct	336
Db	201	CTCACAAAGCCCAATGCTGAGGAGATGACCAAGTACCACACAGCATGACTACATAATTCT	260
OY	337	tgcgctccatccgtccagataaacatgttcggaatcacgcaagcagatgcagagattcaacg	396
Db	261	TGCGCTCCATCCGTCCTCCAGATAACATGTCGGAGTACACGCAAGCAGATGCAGAGATTCACAG	320
OY	397	tttgtgaggaactgtccagtatctcgatggcctgtttgagltctgtcagltgttctactgttg	456
Db	321	TTGGTGAGGACTGTCCAGTATTTCGATGGCCTGTTGAGTCTGTCTAGTTGTCTACTGTGTG	380
OY	457	gttctgtgccaagtgtctgtgaaccttaataagcagcagacgagacatcgccgtgaattggg	516
Db	381	GTTCTGTGGCAAGTGCCTGTGAACCTTAATAACGACGACGAGACATCGCTGTGAATTGGG	440

OY 517 ctggggcctgcaccatgcaagaagtcaggagcatctgtgttcttcaatgata 576
|||||
Db 441 CTGGGGCCCTGCACCATGCACAAGAAGTCCGAGGCATCTGGCTTCTGTTACGTCATGATA 500
OY 577 tcgtcttggccatcctcgtgaactgtctaagtatcaccaagaagtgctgtacattgacattg 636
|||||
Db 501 TCGTCTTGGCCATCTCTGGAAGTCTTAAGTATCACAGAGGGTGTGTACATTGACATTG 560
OY 637 atattcaccatggtgacgagcgtggaagagccctctacaccacgagcgggtcatgactg 696
|||||
Db 561 ATATTCAACCATGCTGACGGCGTGGAGAGGCCCTTCTACACACGACCGGGTCATGACTG 620
OY 697 tgtccttcataaagtatgagagtaactcccaagaaactgggagacctgaagatataccgag 756
|||||
Db 621 TGTCTTTCATTAAGTATGAGAGTACTTCCACGAAGTGGGACCTACGGGATATCGGGG 680
OY 757 ctggcaagaacaagtattatgtgttaactaccgcgtccgagagcggtatgtacagagt 816
|||||
Db 681 CTGGCAAGGCAAGTATATGCTGTAACTTACCGCTCCGAGACGGGATTGATGACCAGT 740
OY 817 cc--tatgagccatttccaagccggtcatgttccaaagtaatgagaatgttcc 867
|||
741 CCTATGAAGGCATTTTCAACCCGGCCATGGCCAAAGTAATGGAGAAGTTCC 793

Search completed: June 8, 2002, 11:39:48
Job time: 7656 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 09:35:07 ; Search time 3028.56 Seconds
(without alignments)
11131.578 Million cell updates/sec

Title: US-09-817-538-2
Perfect score: 1611
Sequence: 1 atgtctgggtctctgccg.....tccctcagcttcttcccc 1611

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1596.6	99.1	1611	6	AX053101	AX053101 Sequence
2	1596.6	99.1	1611	9	HSU50079	U50079 Human histo
3	1499.6	93.1	2093	9	BC000301	BC000301 Homo sapi
4	1498	93.0	2091	9	D50405	D50405 Human mRNA
5	1498	93.0	2111	6	AR012007	AR012007 Sequence
6	1498	93.0	2111	6	162388	162388 Sequence 2
7	1498	93.0	2111	23	E11455	E11455 cDNA encodi
8	1240.2	77.0	1977	10	MMHIDE	X98207 M.musculus
9	1236.4	76.7	1999	10	MMHIDE	U80780 Mus musculu
10	1110.4	68.9	81975	2	AF000450	AP000450 Homo sapi
11	1110.4	68.9	146841	9	AC020549	AC020549 Homo sapi
12	1098.4	68.2	187852	2	AC068041	AC068041 Homo sapi
13	1054	65.4	252957	2	AC105705	AC105705 Rattus no
14	997.4	61.9	1646	5	AF044169	AF044169 Gallus ga
15	997.4	61.9	1678	5	AF039751	AF039751 Gallus ga
16	992.6	61.6	1646	5	AF043328	AF043328 Gallus ga
17	953	59.2	1487	5	AF020658	AF020658 Xenopus l
18	943.4	58.6	2305	5	XLAB21	X78454 X.laavis AB
19	873.6	54.2	3701	5	CHKTCKL	J03579 Chicken tyr
20	873.6	54.2	4073	5	GGTK1	X60380 G.gallus tk
21	821.2	51.0	1878	5	AF039752	AF039752 Gallus ga
22	818	50.8	1997	6	AX305742	AX305742 Sequence
23	818	50.8	1997	10	MMU031758	U31758 Mus musculu
24	812.8	50.5	101767	9	AL389885	AL389885 Human DNA
25	811.2	50.4	140242	2	AC073038	AC073038 Homo sapi
26	805.2	50.0	1985	6	AX053103	AX053103 Sequence
27	805.2	50.0	1985	9	HSU31814	U31814 Human trans
28	752.2	46.7	121601	9	AL592406	AL592406 Human DNA
29	752.2	46.7	154386	2	AC016314	AC016314 Homo sapi
30	752.2	46.7	159286	2	AC025710	AC025710 Homo sapi
31	734.4	45.6	2180	3	AF032919	AF032919 Strongylo
32	729	45.3	2157	3	AY058487	AY058487 Drosophi
33	717.8	44.6	2172	3	DMHISSEAC	Y09258 D.melanogas
34	712.8	44.2	2073	2	AF026949	AF026949 Drosophi
35	655.8	40.7	106840	2	AC094420	AC094420 Rattus no
36	654.6	40.6	70949	2	AC095666	AC095666 Rattus no
37	642.2	39.9	2681	3	AF086715	AF086715 Drosophi
38	640.6	39.8	51246	2	AC014151	AC014151 Drosophi
39	640.6	39.8	156525	3	AC010665	AC010665 Drosophi
40	640.6	39.8	301769	3	AE003482	AE003482 Drosophi
41	611.8	38.0	768	10	AF321129	AF321129 Rattus no
42	583	36.2	45754	9	AL139118	AL139118 Human DNA
43	551.6	34.2	158526	2	AC094458	AC094458 Rattus no
44	544.2	33.8	106731	2	AC097622	AC097622 Rattus no
45	532	33.0	109924	2	AC105535	AC105535 Rattus no

ALIGNMENTS

RESULT 1
LOCUS AX053101 1611 bp DNA
DEFINITION Sequence 25 from Patent WO0071703.
ACCESSION AX053101
VERSION AX053101.1 GI:12227157
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1611)
AUTHORS Macleod,A.R., Li,Z. and Besterman,J.M.
TITLE Inhibition of histone deacetylase
JOURNAL Patent: WO 0071703-A 25 30-NOV-2000;
Methylgene, Inc. (CA)
FEATURES
source location/Qualifiers
1..1611
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 428 a 385 c 440 g 358 t
ORIGIN

Query Match	99.18;	Score 1596.6;	DB 6;	Length 1611;
Best Local Similarity	99.48;	Pred. No. 0;		
Matches 1602; Conservative	0;	Mismatches	9;	Indels 0;
				Gaps 0;

QY	1	atgtctggggtctctgtccccgtgtgtgtctgtctgtctccccactcgtgtcatccctgagaacaca	60
Db	1	ATGTCGTGGGGTCTCTGCCCCCTGGTGTCTGTCTCCACTCGGTTCATCCTGAGAACA	60
QY	61	gcctgaagcgcrcctgttcaactcggggtagaccacgcggggaagcgaagaatgycgcaga	120
Db	61	GCCTGAGCGTCTCTGTCTACTCGGGGTAGACCACGCGGGGAGCGAGCAAGATGCGCGAGA	120
QY	121	cgaagcgaccccgaggaagaagtcgttactactacgacgggagtglttgaaattactatc	180
Db	121	CGCAGGGCACCCCGAGGAAAGTCTGTTACTACTACGACGGGATGTTGAAATTAACFAAT	180
QY	181	atgacacaagggccaccaatgaagcctcacccaatccgcatgactcataatttgtctgtca	240
Db	181	ATGGACAAGGCCACCCCAATGAAGCCTCACCGAATCCGCAATGACTCATATTTGCTGTCA	240
QY	241	actatggtctctaccgnaaaatggaatctatcgccctcacaaagccaatgtctgagaga	300
Db	241	ACTATGCTCTCTACCGAAATAATGAATCTATCGCCCTPCACAAGCCATGCTGAGGAGA	300
QY	301	tgaccaagtlaccacagcgatgactacattaaattcttgcgtccatccgtccagataaca	360
Db	301	TGACCAAGTACCACAGCGATGACTACATTAATTCCTGGCGCTCCATCCGTCCAGATACAA	360
QY	361	tgtcgagtagtacgcaagcagatgacagagattcaacgttggtaggactgtccagtatccg	420
Db	361	TGTCGGAGTACAGCAGCAGATGCAGAGATTCAACGTTGGTGAGGACTGTCCAGTATTCG	420
QY	421	atggcctgtttgagttcctgtcagttgtcctactgtgtgtctgttggcaagtgtgtgaac	480
Db	421	ATGGCCTGTTGAGTTCGTCTCAGTTGTCTACTGTGTGTCTGTGGCAAGTGCTGTGAAC	480
QY	481	ttaataagcagcagacgagacatcgcgctgaattggcgctggggcctgacccatgcaaga	540
Db	481	TTAATAAGCAGCAGACGACATCGCTGTGAATTGGCGTGGGGCCTGCACCATGCACAAGA	540
QY	541	agtcgagagcatctggtctctgttcactcaatgatatcgcttggccatccctgnaactgc	600
Db	541	AGTCGAGGACATCTGGCTTCTGTTACGTCATGATATCGTCTTGGCCATCCTGGAACCTGC	600
QY	601	taagatatcacccagaggggtgctgtacattgacattgatattacccaatggtgaacgctgg	660
Db	601	TAAAGTATCACCCAGAGGGTCTGTACATTGACATTGATATTCACCATGGTGACCGCGGTGG	660
QY	661	aagagggcctctacacccaacgagccgggtcatgactgtgtcccttcataaaglatgagagt	720
Db	661	AAGAGGGCTTCTACACCACGGAACCGGGTCATGACTGTGTCTTTCATAAGTATGAGAGT	720
QY	721	acttcccaagaaactggggaacctacgagataccggggctgycaaaagacaagtatatgctg	780
Db	721	ACTTCCCAAGAACTGGGGAACCTACGGGATATCGGGGTGGCAAAAGCAAGTATTATGCTG	780
QY	781	ttaactacccgctccgagacgggalttgatgacgagttcctatgagccatttcagaacgg	840
Db	781	TTAACTACCCGCTCCGAGACGGGATTGATGACGAGTCTCTATGAGGCCATTTCAGACCGG	840
QY	841	tcaatgccaaaagtaatgagatgttccagcctagtgcggtgtcttacagtgtygtcag	900
Db	841	TCATGTCCAAAGTAATGAGATGTTCCAGCCCTAGTGGGTGCTTTACAGTGTGGCTCAG	900
QY	901	actcccatctggggatcgggttaggttgcttcaatctatctatacaaaagacacgcceaagt	960
Db	901	ACTCCCTATCTGGGGATCGGGTTAGGTTGCTTCAATCTAATCTAATCAAAAGACACGCCCAGT	960
QY	961	gtgtggaatttgtcaagaagctttaacctgcctatgtctgtagctgggagcggtgtgtaca	1020
Db	961	GTGTGGAATTTGTCAAGACCTTTAACTTGCCCTATGCTGATGCTGGGAGCGGTGTGTACA	1020

QY	1021	ccattcgtaacgcttgcgccggtgctgagacatatgagacagctgtgcccctgagatacgaga	1080
Db	1021	ccattcgtaacgcttgcgccggtgctgagacatatgagacagctgtgcccctgagatacgaga	1080
QY	1081	ttccctaatagacttccataaatgactactttgaataactttgagccagatttcaagctcc	1140
Db	1081	ttccctaatagacttccataaatgactactttgaataactttgagccagatttcaagctcc	1140
QY	1141	acatcagtccttccaatatgactaacccagaaacggaatgagtacctgtgagaagaatcaaac	1200
Db	1141	acatcagtccttccaatatgactaacccagaaacggaatgagtacctgtgagaagaatcaaac	1200
QY	1201	agcgactgtttgagaaccttagaatgtctgcgcgacgcacctgggtccaaacgcgacgcga	1260
Db	1201	agcgactgtttgagaaccttagaatgtctgcgcgacgcacctgggtccaaacgcgacgcga	1260
QY	1261	ttccttgagagcgcaccccttgagagagatgycgatgagagcagaagacgaccttgacaagc	1320
Db	1261	ttccttgagagcgcaccccttgagagagatgycgatgagagcagaagacgaccttgacaagc	1320
QY	1321	gcattctgatactgtcctcctctgcacaanaogaaattgacctgtgaggaagagttctccgattctg	1380
Db	1321	gcattctgatactgtcctcctctgcacaanaogaaattgacctgtgaggaagagttctccgattctg	1380
QY	1381	aagagggagggagaggggggcccgaagaactcttccaacttcaaaaaaagccaagagagtca	1440
Db	1381	aagagggagggagaggggggcccgaagaactcttccaacttcaaaaaaagccaagagagtca	1440
QY	1441	aaacagaggaatgaaaaagagaaaagacccagagagagaagaagaatcaaccgaagagaga	1500
Db	1441	aaacagaggaatgaaaaagagaaaagacccagagagagaagaagaatcaaccgaagagaga	1500
QY	1501	aaaccaagagagagaagccagaagccaaggggtcaagagagagggccaagttgacctgaa	1560
Db	1501	aaaccaagagagagaagccagaagccaaggggtcaagagagagggccaagttgacctgaa	1560
QY	1561	tggacctctccagctctggtcttccctgtgagtcacctcaagtttcttcccc	1611
Db	1561	tggacctctccagctctggtcttccctgtgagtcacctcaagtttcttcccc	1611

RESULT	2
LOCUS	HSU50079 1611 bp mRNA linear PRI 16-MAY-1996
DEFINITION	Human histone deacetylase HD1 mRNA, complete cds.
ACCESSION	U50079
VERSION	U50079.1 GI:1277083
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 1611)
JOURNAL	Taunton, J., Hassig, C.A. and Schreiber, S.L.
MEDLINE	A mammalian histone deacetylase related to the yeast
REFERENCE	transcriptional regulator Rpd3p
AUTHORS	Science 272 (5260), 408-411 (1996)
JOURNAL	96185499
MEDLINE	2 (bases 1 to 1611)
REFERENCE	Taunton, J., Hassig, C.A. and Schreiber, S.L.
AUTHORS	Direct Submission
JOURNAL	Submitted (27-FEB-1996) Jack Taunton, Chemistry, Harvard
FEATURES	University, 12 Oxford, Cambridge, MA 02138, usa
SOURCE	Location/Qualifiers
	1..1611
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/cell_type="Jurkat T cell"
	111..1559
	/note="similar to S. cerevisiae RPD3, a global
	transcriptional regulator; trapoxin receptor"
	/codon_start=1
	/product="histone deacetylase HD1"

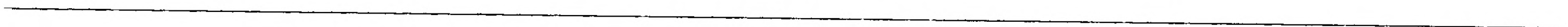
✓ member
phobos
pinner



Query Match	99.18;	Score 1596.6;	DB 9;	Length 1611;
Best Local Similarity	99.48;	Pred. NO. 0;		
Matches 1602; Conservative	0;	Mismatches	9;	Indels 0;
		Gaps		0;

RESULT 3
BC000301 LOCUS BC000301 2093 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, histone deacetylase 1, clone MGC:8378 IMAGE:2820260,
mRNA, complete cds.
ACCESSION BC000301
VERSION BC000301.1 GI:12653070
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2093)
Strausberg, R.
Direct Submisston
Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

○



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 09:35:07 ; Search time 3028.56 Seconds
(without alignments)
11131.578 Million cell updates/sec

Title: US-09-817-538-2
Perfect score: 1611
Sequence: 1 atgtctgggtctctctgccg.....tcctcagttcttccccc 1611

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBankl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Query Score Match Length DB ID Description

1	1596.6	99.1	1611	6	AX053101
2	1596.6	99.1	1611	9	HSU50079
3	1499.6	93.1	2093	9	BC000301
4	1498	93.0	2091	9	D50405
5	1498	93.0	2111	6	AR012007
6	1498	93.0	2111	6	I62388
7	1498	93.0	2111	23	E11455
8	1240.2	76.7	1997	10	MMHIDE
9	1236.4	76.7	1999	10	MMHIDE
10	1110.4	68.9	81975	2	AP000450
11	1110.4	68.9	146841	9	AC020549
12	1098.4	68.2	187852	2	AC068041
13	1054	65.4	252957	2	AC105705
14	997.4	61.9	1646	5	AF044169
15	997.4	61.9	1678	5	AF039751
16	992.6	61.6	1646	5	AF043328
17	953	59.2	1487	5	AF020658
18	943.4	58.6	2305	5	XLAB21
19	873.6	54.2	3701	5	CHKTCKL
20	873.6	54.2	4073	5	GGTK1
21	821.2	51.0	1878	5	AF039752
22	818	50.8	1997	6	AX305742
23	818	50.8	1997	10	MMU31758
24	812.8	50.5	101767	9	AL389885
25	811.2	50.4	140242	2	AC073038
26	805.2	50.0	1985	6	AX053103
27	805.2	50.0	1985	9	HSU31814
28	752.2	46.7	121601	9	AL592406
29	752.2	46.7	154386	2	AC016314
30	752.2	46.7	159286	2	AC025710
31	734.4	45.6	2180	3	AF032919
32	729	45.3	2157	3	AY058487
33	717.8	44.6	2172	3	DMHISSEAC
34	712.8	44.2	2073	3	AF026949
35	655.8	40.7	106840	2	AC094420
36	654.6	40.6	70949	2	AC095666
37	642.2	39.9	2681	3	AF086715
38	640.6	39.8	51246	2	AC014151
39	640.6	39.8	156525	3	AC010665
40	640.6	39.8	301769	3	AE003482
41	611.8	38.0	768	10	AF321129
42	583	36.2	45754	9	AL139118
43	551.6	34.2	158526	2	AC094458
44	544.2	33.8	106731	2	AC097622
45	532	33.0	109924	2	AC105535

ALIGNMENTS

AX053101 Sequence
U50079 Human hist
BC000301 Homo sapi
D50405 Human mRNA
AR012007 Sequence
I62388 Sequence 2
E11455 cDNA encod
X98207 M.musculus
U80780 Mus musculu
AP000450 Homo sapi
AC020549 Homo sapi
AC068041 Homo sapi
AC105705 Rattus no
AF044169 Gallus ga
AF039751 Gallus ga
AF043328 Gallus ga
AF020658 Xenopus l
X78454 X.laavis AB
J03579 Chicken tyr
X60380 G.gallus tk
AF039752 Gallus ga
AX305742 Sequence
U31758 Mus musculu
AL389885 Human DNA
AC073038 Homo sapi
AX053103 Sequence
U31814 Human trans
AL592406 Human DNA
AC016314 Homo sapi
AC025710 Homo sapi
AF032919 Strongylo
AY058487 Drosophi
Y09258 D.melanogas
AF026949 Drosophi
AC094420 Rattus no
AC095666 Rattus no
AF086715 Drosophi
AC014151 Drosophi
AC010665 Drosophi
AE003482 Drosophi
AF321129 Rattus no
AL139118 Human DNA
AC094458 Rattus no
AC097622 Rattus no
AC105535 Rattus no

Linear PAT 12-JAN-2001

RESULT 1
AX053101
LOCUS AX053101 1611 bp DNA
DEFINITION Sequence 25 from Patent WO0071703.
ACCESSION AX053101
VERSION AX053101.1 GI:12227157

SOURCE

ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Macleod,A.R., Li,Z. and Besterman,J.M.
TITLE Inhibition of histone deacetylase
JOURNAL Patent: WO 0071703-A 25 30-NOV-2000;
Methylene, Inc. (CA)

FEATURES
Source Location/Qualifiers
1..1611

BASE COUNT 428 a 385 c 440 g 358 t
ORIGIN /organism="Homo sapiens"
/db_xref="taxon:9606"

Query Match	99.18;	Score 1596.6;	DB 6;	Length 1611;
Best Local Similarity	99.48;	Pred. No. 0;		
Matches 1602; Conservative	0;	Mismatches 9;	Indels 0;	Gaps 0;

QY	1	atgtctgggggtctctgtccccctgtgtgtctgtctgtctccactcgtgtcatcctgagaacaca	60
Db	1	ATGTCGTGGGGTCTCTGTCCCCGTGTGTGCTGTGTCTCCACTCGGTTCATCTTGAGAACACA	60
QY	61	gccttagcgrrctctgtcactcggggtagaccacgcgggaggcgagcaagatgycgcaga	120
Db	61	GCCTGAGCGTCTCTGTCTACTCGGGGTAGACCACGCGGGGAGCGAGCAAGATGGCGCAGA	120
QY	121	cgcagggcacccgcgaggaagtctgttactactacgacgggagtgttgaataattactatc	180
Db	121	CGCAGGGCACCCGAGAGAAAGTCTGTACTACTACGACGGGAGTGTGGAATTTACTATT	180
QY	181	atgacacaagccaccacaatgtgaagccctcacccgaatccgcatgactcataatttgcgtcca	240
b	181	ATGGACAAGGCCACCCCAATGAAGCCCTCACCGAATCCGCATGACTCATATAATTTGCTGTCA	240
QY	241	actatggtctctaccgaaaaatgtgaatctatcgcctccacaaagccaatgtctgagaga	300
Db	241	ACTATGCTCTCTACCGAAATAATGGAATCTATCGCCCTCCACAAGCCAATGCTGAGGAGA	300
QY	301	tgaccaagtaccacacagcgaatgactacattaaattcttgcgtccatccgtccagataaca	360
Db	301	TGACCAAGTACCACACAGCGATGACTACATTAAATTTCTGGCTCCATCCGTCCAGATATACA	360
QY	361	tgtcggagtacagcaagcagaatgcagagattcaacgttggtaggactgtccaglattcg	420
Db	361	TGTCGGAGTACAGCAAGCAGATGCAGAGATTC AACGTTGGTGAGGACTGTCTCAGATATTCG	420
QY	421	atggcctgtttgaagttcgttcagttgttctactgtgtgtctgttggcaagtgtctgaac	480
Db	421	ATGGCCTGTTGAGTTCGTCTGAGTTGTCTACTGGTGTGTCTGTGGCAAGTGCTGTGAAC	480
QY	481	ttaataagcagcagacgagacatcgccgtgaattggcgttggggccttgcaccatgcaaga	540
Db	481	TTAATAAGCAGCAGACGAGACATCGCTGTGAATTTGGGCTGGCGCCTGCACCATGCACAAGA	540
QY	541	agtcgagagcatctggtctctgttcaagtcacatgatatcgtcttggccatccctgnaactgc	600
Db	541	AGTCCGAGGCATCTGGCTTCTGTACGTCAATGATATCTCTTGGCCATCCTGGAACCTGC	600
QY	601	taaaagtataccacagaggtgtgttacatttgacattgataattcaccaatgttgacgctg	660
Db	601	TAAAGTATCACCCAGAGGGTGTGTACATTGACATTGATATTCAACCATGTTGACGGCGGTGG	660
QY	661	aagagcgcttctacacacacgagaccgggtcatgactgtgtcccttcatagaatgagagat	720
Db	661	AAGAGGGCTTCTACACCCAGGACCGGGGTCATGACTGTGTCTTTCATATAAGTATGAGAGT	720
QY	721	acttcccaagaaactggggaaccttcaagygataaccggggcctgycaaagaacagtattatctg	780
Db	721	ACTTCCCAAGAACTGGGACCTACGGGATATCGGGGCTTGGCAAAAGGCAAGTATTATGCTG	780
QY	781	ttaactacccgctccgagacggaatttgatgacgaagttcctatgagccatttcaagccgg	840
Db	781	TTAACTACCCGCTCCGAGACGGGATTTGATGACGAGTCTATGAGGCCATTTTCAAGCCGG	840
QY	841	tcatgtccaaagtaatgagagatgttccagcctagtgcggtgtcttacagttgtgtccag	900
Db	841	TCATGTCCAAAGTAATGAGAGTGTCCAGCCCTAGTGGCGGTGCTTTACAGTGTGGCTCAG	900
QY	901	actcccatctggggatcggttaggttgccttcaatctatctatacaaaagacacgcceaagt	960
Db	901	ACTCCCTATCTGGGGATCGGTTAGGTGTGCTTCAATCTAATCAATCAAAAGGACACGCCCAAGT	960
QY	961	gtgttgaaattgtcaaaagcittaaacctgcctatgtctgatgtctgggagcggtgttaca	1020
Db	961	GTGTGGAATTTGTCAAGAGCTTTAACTGCCTATGCTGATGCTGGAGGCGGTGTGTTACA	1020

QY	1021	ccattcgtaacggtgcccgggtgctgagacatatgagacagctgtgyccttgatcggaga	1080
Db	1021	ccattcgtaacggtgcccgggtgctgagacatatgagacagctgtgyccttgatcggaga	1080
QY	1081	tcacctaatgagcttcatacaatgactactttgaaactttgaccagatttcaagctcc	1140
Db	1081	tcacctaatgagcttcatacaatgactactttgaaactttgaccagatttcaagctcc	1140
QY	1141	acatcagtccttccaatatgactaaaccagaacacgaatgagtacctgagagaatcaaac	1200
Db	1141	acatcagtccttccaatatgactaaaccagaacacgaatgagtacctgagagaatcaaac	1200
QY	1201	agcgactgtttgagaaccttagaatgtgtccgcacgcaccttgygtccaacgcagcgca	1260
Db	1201	agcgactgtttgagaaccttagaatgtgtccgcacgcaccttgygtccaacgcagcgca	1260
QY	1261	ttcctgagagcgcacccctgagagagatgtagcatgagagcagaacgcaccttgacaagc	1320
Db	1261	ttcctgagagcgcacccctgagagagatgtagcatgagagcagaacgcaccttgacaagc	1320
QY	1321	gcattctcgatctgtctctctctgtacaacaacgaattgtccctgtaggaagagttctcagattctg	1380
Db	1321	gcattctcgatctgtctctctctgtacaacaacgaattgtccctgtaggaagagttctcagattctg	1380
QY	1381	aagagagagagagagggggccgcgaagaactcttccaacttcaaaaaagccaagaagagtc	1440
Db	1381	aagagagagagagagggggccgcgaagaactcttccaacttcaaaaaagccaagaagagtc	1440
QY	1441	aaacagagatgaaaaaagaagaacccagagagagaagaagaatcacccgaagaaggaga	1500
Db	1441	aaacagagatgaaaaaagaagaacccagagagagaagaagaatcacccgaagaaggaga	1500
QY	1501	aaaccaagagagagaaagccagaagccaagaggtcaagagagagagccaagttgagctgaa	1560
Db	1501	aaaccaagagagagaaagccagaagccaagaggtcaagttgagctgaa	1560
QY	1561	tggacctctccagctcttggcttctctgtgtgagtcacctcaagtttcttccccc	1611
Db	1561	tggacctctccagctcttggcttctctgtgtgagtcacctcaagtttcttccccc	1611

RESULT 2
HSU50079 LOCUS 1611 bp mRNA linear PRI 16-MAY-1996
DEFINITION Human histone deacetylase HD1 mRNA, complete cds.
VERSION U50079
KEYWORDS U50079.1 GI:1277083
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1611)
TITLE Taunton, J., Hassig, C.A. and Schreiber, S.L.
JOURNAL A mammalian histone deacetylase related to the yeast
MEDLINE transcriptional regulator Rpd3p
REFERENCE Science 272 (5260), 408-411 (1996)
AUTHORS 2 (bases 1 to 1611)
TITLE Taunton, J., Hassig, C.A. and Schreiber, S.L.
JOURNAL Direct Submission
TITLE Submitted (27-FEB-1996) Jack Taunton, Chemistry, Harvard
JOURNAL University, 12 Oxford, Cambridge, MA 02138, usa
FEATURES
source Location/Qualifiers
1.1611
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="Jurkat T cell"
111.1559
/note="similar to S. cerevisiae RPD3, a global
transcriptional regulator; trapoxin receptor"
/codon_start=1
/product="histone deacetylase HD1"

```

/protein_id="AAC50475.1"
/db_xref="GI:1277084"
/translation="MAQTQGTTRRKVCYYDGDVGNYYYGQGHMPKPHRIEMTHNLLN
YGLYRKMEITYRPHKANAEEMTKYHSDYIKFLRSIRPDNMSEYSKQMRNVGDCRPV
FDGLFEFCQLSTGGSVASAVKLNKQOTDIAVNWAGGLHNAKKSSEASGFCVNDIYLA
LELKYHQRVLYIDIDIDIHGDGVEEAFYTDRTVMTVSFHKGYEYFPETGDLRIDGAGK
GKYIAVNPPLRDGIDIDESYEAIKFPMSRVMEFMQPSAVVLQCGSDSLGDLGCFNL
TIKHAKEVEFKSFENLPMIMGGGYTIIRNVARCWTYETAVALDTEIPNELPYNDYF
EYFGPDKLHISPSNMTNONTNEYLEKIQLFENLRMLPHAPGVQMAIPEDAIPEE
SGDEDEDPPDKRISICSSDKRIACEEESDSEEGEGGRKNSSNFKAARVKTDEKE
KDPEEKKEVTEEEKTEKEEPEAKGVKEEVKLA"

```

Query Match	99.18;	Score 1596.6;	DB 9;	Length 1611;
Best Local Similarity	99.48;	Pred. No. 0;		
Matches 1602; Conservative	0;	Mismatches 9;	Indels 0;	Gaps 0;

Qy 1 atgtctgggtctctcgcccgctgtgtctgtctctccactcgttcattcctgagaacaca 60
|||||
1 ATGTCTGGGGTCTCTGCCCGCTGTGTCTGTCTCTCTCTCCACTCGGTATCTCTGAGAACACA 60

Qy 61 gacctagagcrtctgttactcggggttagaccacgcggggagcgagcaagatfygcaga 120
|||||
Db 61 GCGTGAAGCTCTCTGTCACTCGGGGTAGACCACCGCGGAGGCGAGCAAGATGCGCAGA 120

Qy	121	CGCAGGACCCGAGGAAGTCTGTACTACTACGACGGATGTTGGAATTACTATT	180
Db	121	CGCAGGACCCGAGGAAGTCTGTACTACTACGACGGATGTTGGAATTACTATT	180

```
OY      181 atgacaaagccaccaatgaagcctcacccaatccgcataactcataattgtcgctca 240
        |||||
Db       181 ATGACAAGGCCACCCAAATGAAGCCTCACCGAATCCGCATGACTCATTAATTGTGCTCA 240
```

QY 241 actatggtctctaccgaaaaatgtaaatctatcgccctcacaagaagccaatgctgaagaga 300
|||||
Db 241 ACTATGCTCTCTACCGAAAAATGGAATCTATCGCCCTCACAAAGCCAATGCTGAGGAGA 300

[illegible]

QY	361	tgtcgga	gtacagca	agcagat	gtcaga	gattca	acgltgt	gtga	gactgt	ccagtatt	cg	420
Db	361	tgtcgga	gtacagca	agcagat	gtcaga	gattca	acgltgt	gtga	gactgt	ccagtatt	cg	420

QY 421 atggcctgtttgagttctgtcagttgtctactgtgtgttcgtlgtgcaagtctgttgaac 480
|||||
DB 421 ATGGCCTGTTGAGTCTGTGCAAGTGTCTACTGTGGTGGTCTGTGTGCAAGTGCCTGTAAC 480

QY	481	ttaataagcagcagacgcgctgaaattggcgctggggccgtcaccatycaaaga	540
Db	481	TTAATAAGCAGCAGACGCATCGCTGGAATTGGCGTGGGGCCTCACCATGCAAGA	540

QY 541 agtcgaggcacctcgtgctctgtttacgtcaatgatatactgcttggccacctctygaactgc 600
|||||
541 agtccgaggcacctcgtgctctgtttacgtcaatgatatactgcttggccacctctygaactgc 600
|||||

QY	601	taaagtatcaccagagggtgctgtacattgacattgatataccaacatggttaacygccttgg	660
Db	601	TAAGTATCACCAGAGGGTGTGTCATTGCACATTGATATTCACCCATGGTGACCGCGCTGG	660

Qy 661 aagagcctctacaccacgagcggtcatgactgtgtcccttcataatgatgtgagagt 720
|||||
|||
Db 661 AAGAGGCGCTTCTACACCACGAGCGGTCATGACTGTGTCCTTCATTAAGTAAGGAGAGT 720

```

QY      721  acttccaggaaactggygacctacgagataccgggctgycgaaagacaagatattatgctg 780
        |||||
Db      721  acttccaggaaactggggacctacgggatatcggggctggcaaaaggcaagtattatgctg 780

```

QY 781 ttaactaccgcgtccgagacgagatgatgacgagtlcctatgtagggccatttccaagccg 840
|||||

Db	781	TTAACTACCCGCTCCGAGACGGGATGATGACGAGTCTCTATGAGGCCATTTCGAAGCCGG	840
QY	841	tcatactccaaaglaatgagatgttccagccctagtcgctgtgtcttaacagtgtgctcag	900
Db	841	TCATGTCCAAAGTAATGGAGATGTTCCAGCCTAGTGCGGTGGTCTTACAGTGTGGCTCAG	900
QY	901	actccatatctggggatcgggttagtgttgcctcaatctatctatccaagagcacgccaagt	960
Db	901	ACTCCCTATCTGGGGATCGGTTAGGTGCTTCAATCTAATCTATCAAGGACACGCCCAAGT	960
QY	961	gtgtggaatttgtcaagagccttaacctgctatgtctgatgtctgtaggctgtgttaca	1020
Db	961	GTTGTGAATTGTGCAAGAGCTTTAACTTGCTCTATGCTGATGCTGGAGCGGTGTTACA	1020
QY	1021	ccatctgtaacglttgcccggtgtctgacatalgagacagctgtgcccctgatacggaga	1080
Db	1021	CCATCTGTAACGTTGCCCGGTGCTGGACATATGAGACAGCTGTGGCCCTGGATACGGAGA	1080
QY	1081	tccttaatgagcttccatacaatgtaactttgaaacttltggaaccaatttcaagctcc	1140
Db	1081	TCCCTTAATGAGCTTCCATACAAATGACTACTTTGAATACTTTGGACCAGATTTCAGCTCC	1140
QY	1141	acatcagtccttccaatatgactaaccagaacaacggaatgagtaacctgtagaagaatcaac	1200
Db	1141	ACATCAGTCTCTTCCAATATGACTAACCCAGAACACGAATGAGTACTTGAGAGAAAGATCAAAC	1200
QY	1201	agcgactgtttgagaaccttagaatgtctgcgcgcacgcacactgtggttccaacgcagcgca	1260
Db	1201	AGCGACTGTTTGAGAACCTTAGAATGCTGCCGCACGCACCTGGGGTCCAATAATGCAGGCCGA	1260
QY	1261	ttcctgagagcgcacatcccctgagagagtggcgatgagacgaagacgacccctgacaagc	1320
Db	1261	TTCCTGAGGACGCCCATCCCTGAGGAGAGTGGCGATGAGGACGAAGACGACCCCTGACAAGC	1320
QY	1321	gcattctgatctgtctcctctgacaaaacgaatttgcctgtgaggaagagtcttccgattctg	1380
Db	1321	GCATCTGATCTGTCTCTGTGACAAACGAATTGCCCTGTGAGGAAGAGTTCCTCCGATTCTG	1380
QY	1381	aagaggaaggagaggggggcgcgaagaactcttccaacttcaaaaaagccaagagagtca	1440
Db	1381	AAGAGGAGGGAGAGGGGGCGCAAGACTCTTCCAACTTCAAAAAAGCCAAAGAGAGTCA	1440
QY	1441	aaacagaagatgaaaaagagaaagaccagagagagaagaagaatcacccgaagagaga	1500
Db	1441	AAACAGAGGATGAAAAAGAGAAAGACCAGAGAGAGAAGAAAGAAAGATCACCCGAAGAGAGA	1500
QY	1501	aaaccaagagagagaagccagaagccaagggttcaagagagagcgaagtgtgacctgaa	1560
Db	1501	AAACCAAGAGAGAGAAAGCCAGAACCAAGGGGTCAAGAGAGGAGGTCAAGTTGGCCTGAA	1560
QY	1561	tggacctctccagctctgtgcttccgtgtagttccctcaacgttcttctccc	1611
Db	1561	TGGACCTCTCCAGCTCTGTGGCTTCTGCTGAGTCCCTCACGTTTCTTCCCC	1611

RESULT	3
BC000301	
LOCUS	BC000301
DEFINITION	Homo sapiens, histone deacetylase 1, clone MGC:8378 IMAGE:2820260, mRNA, complete cds.
	2093 bp mRNA linear PRI 12-JUL-2001

ACCESSION	BC000301
VERSION	BC000301.1
	GI:12653070

KEYWORDS	MGC.
SOURCE	human.

ORGANISM	Homo sapiens
Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:	

REFERENCE
1 (bases 1 to 2093)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS	Strausberg, R.
TITLE	Direct Submission

JOURNAL Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC) Cancer Genomics Office National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK
COMMENT

USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/nisc_mgc/nhgrl.nih.gov
Contact: nisc_mgc@nhgrl.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Masieilo, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 1 Row: c Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13128859.
Location/Qualifiers

FEATURES
source

1. .2093
/organism="Homo sapiens"
/db_xref="locusID:15181"
/db_xref="taxon:9606"
/clone="MGC:8378 IMAGE:2820260"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NIH_MGC_7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
44. .1492
/codon_start=1
/product="histone deacetylase 1"
/protein_id="AAH00301.1"
/db_xref="gi:12653071"
/translation="MAQTGTRKVCYYVDGVNYYVGQHPMKPHRTMTHNLLN
YGLRKMEIYRPHKANAEMTKYHSDYIKFLRSIRPDNMSEYSKOMRPNVGEDCPV
FDGLFEFCQJSTGGSVASAVKLNKQDTIAVNMAGGLHHAKSEASGFCYVNDIVLAI
LELKYHORVLYIDIDIHGDVEEAFYTRVMTVSFHKYGEYFPGCDLRDIGAGK
GKYAVNYPRLDGDIDDESYEAIKPKVMKVMEMFQPSAVVLQCGSDLSGDRIGCFNL
TIKGAKCYEFYKSEFNLPLMLTGGGYTRNVARCWYETELVALDTEIPNLYNDYF
EYFGPDFKLHISPSNMTNONTNEYLEKIKORLFENLRMLPHAPGVQMAIPEDAIPEE
SGDEDEDPRKRISICSSDKRIACEEESDSEEEGEGGRKNSSNFKAKRKYKTEDEKE
KDPEEKKEVTEEEKTKKEKPEAKGVKEEVKIA"

CDS

BASE COUNT 570 a 480 c 531 g 512 t
ORIGIN

Query Match 93.1%; Score 1499.6; DB 9; Length 2093;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 97 gggagcgagcaagatgcgcaagcgaagcgacccggaagaaagtctgtactactacg 156
|||||
Db 30 GGGAGGCGAGCAAGATGGCGAGACGCGAGGCGACCGCGAGAAAGTCTGTACTACTACG 89
QY 157 acggggatgtggaatactatattatgacaagcgccaccaatgaagcctcaccgaatcc 216
|||||
Db 90 ACGGGAGTGTGGAATTACTATTATGACACAGGCCACCAATGAAGCCTCACCAGATTC 149
QY 217 gcatgactcataatttctgtctcaactatgtctctaccgaaaaatgtgaatctatgcc 276
|||||
Db 150 GCATGACTCATATTGCTGCTCAACATATGCTCTTACCAGAAAAATGGAATCTATCGCC 209
QY 277 ctcaacaagccaatgtctgagagagatgaaccaagtlaccacagcgatgactacataaattct 336
|||||
Db 210 CTCACAAAGCCAATGCTGAGGAGATGACCAAGTACCACAGCGATGACTACATTAAATTCT 269

QY 337 tgcgtccatccgctccagatacatgtcggagtagcagcaagcagatgcagagattcaacg 396
|||||
Db 270 TGGCTTCATCCGTCAGATATACTGTCGGAGTACAGCAAGCAGATGCAGAGATTCAACG 329
QY 397 ttggtgaagactgtccagatattcgatgacctgtttgaattctgtcaagttgtctactgtg 456
|||||
Db 330 TTGGTGAGGACTGTCCAGTATTCGATGCGCTGTTTGAATTCTGTCAAGTTGTCTACTGCTG 389
QY 457 gtctgtggaagtgtgtgaaacttaataagcagcagcagacatcgccgtgaattggg 516
|||||
Db 390 GTTCTGTGGCAAGTGTGTAACCTTAATAAGCAGCAGACGACATCGCTGTGAATTGGG 449
QY 517 ctggggcctgcaccatgcgaagaagtcgcgagcattctgtctgttacgtcaatgata 576
|||||
Db 450 CTGGGGCCCTGCACCATGCAGAAGAAGTCCGAGGCATCTGGCTTCTGTACGTAATGATA 509
QY 577 tcgtcttgccatcctgtgaactgtctaagatatacaccagaggggtgtgtacattgacattg 636
|||||
Db 510 TCGTCTTGCCCATCTCGAAGTCTGTAAGTATCACCAGAGGGGTGCTGTACATTGACATTG 569
QY 637 atattcaccatgtgtgacggtgtggaagaggcctctacaccagcagcggtcatcactg 696
|||||
Db 570 ATATTCAACCATGTGTGACGGCGGTGGAAGAGGCCCTTCTACACCAAGGACCGGTCATGACTG 629
QY 697 tgccttcatlaagatagagagtagtctccacgaagactgggagacctacagagataccgggg 756
|||||
Db 630 TGTCTTTTCAATAAGTATGAGAGTACTTCCAGGAAGTGGGACCTACGGGATATCGGGG 689
QY 757 ctggcaaaagacaagatattatgctgttaactaccgcgtccgagacggagattgtagcagagt 816
|||||
Db 690 CTGGCAAAAGCCAAGTATTATGCTGTAACTACCCGCTCCGAGACGGGATTGATGACGAGT 749
QY 817 ccataagagccatttcaagccggtcatgtgccaaagtaatgagatgttccagcctaagt 876
|||||
Db 750 CCTATGAGGCCATTTTCAAGCCGGTCACTGTCCAAAGTAATGAGATGTTCCAGCCTTACTG 809
QY 877 cgtgtgtcttacagtggtgtctcagactccctatctgtggatcgtgttagtcttcaatc 936
|||||
Db 810 CGGTGCTTTACAGTGTGGCTCAGACTCCCTATCTGGGATCGGTTAGTGTCTTCAATC 869
QY 937 tatctatcaaaagcagcagcccaagtgtgtgaaattgttcaagaagcttlaacctgcctatgc 996
|||||
Db 870 TAACTATCAAAAGGACACGCCCAAGTGTGTGGAATTTGTCAAGAGCTTTAACCTGECTATGC 929
QY 997 tgatgtctggagggcgtgtgttaacacattcgtlaacgttgcgggtgtctggaacatagaga 1056
|||||
Db 930 TGATGCTGGAGCGGCTGTTACACCACTTCTGTAACGTTGCCGGTCTGACATATGAGA 989
QY 1057 cagctgtgcccctgatacagagatccctaattgagcttccatacaatgactacttgaat 1116
|||||
Db 990 CAGCTGTGGCCCTGGATACGGAGATCCCTAATGAGCTTCCATATCAATGACTACTTGAAT 1049
QY 1117 actttgaccagatttcaagctccacatcagctccttccataatgactaaccagaacacga 1176
|||||
Db 1050 ACTTTGAGCCAGATTTCAAGCTCCACATCAGTCTTCCAAATATGACTAACCAGAACACGA 1109
QY 1177 atgagtacctggaagaatcaaaacagcgaactgtttgagaaccttagaatgtctgcgcacg 1236
|||||
Db 1110 ATGAGTACCTGGAGAAGATCAAAACAGCACTGTTGAGAACCTTAGAATGCTGCCGACG 1169
QY 1237 caacctgggtccaacagcagcgatctcctgaagagcgccatccctgaggaagatggcgatg 1296
|||||
Db 1170 CACTTGGGTCCAAATGACGCGAFTTCTTGAGGACGCCATCCCTGAGAGAGATGGCGATG 1229
QY 1297 aggacgaagcagaccctgacaaagcgcatctcgatctgtctccttgacaaacgaattgcct 1356
|||||
Db 1230 AGGACGAAGACGACCCGTGACAAAGCGCATCTCGATCTGCTCTGACAAACGAATTGCCCT 1289
QY 1357 gftgaggaagagttctccgattctgaagagagagggagggggccgcaagaactcttcca 1416
|||||
Db 1290 GTGAGGAAGAGTCTCCGATTCTGAAGAGGAGGAGGAGGGGGCCGCAAGAACTCTTCCA 1349
QY 1417 acttcaaaaagccaagagagtgcaaaaacagagagatgaaaaagaagaaccagaggaga 1476

|||||
Db 1350 ACTTCAAAAAGCCAGAGAGTCAAAAACAGAGATGAAAAAGAGAAAGACCAGAGAGAGA 1409
QY 1477 agaaaggaatcaccgaagagagaaaccaaagagagagaaagcagaagcgaaggtca 1536
|||||
Db 1410 AGAAGAAGTCACCCAGAGAGAGAAACCAAGAGAGAGAGAGAGAGAGAGAGGTCA 1469
QY 1537 aggagagagcgaagtgtgctgaatgagacctccagctctgtcttctgtgagtcct 1596
|||||
Db 1470 AGGAGAGAGTCAGTGGCCTGAATGAGACCTCTCCAGCTCTGGCTTCTGCTGAGTCCCT 1529
QY 1597 cagcttcttctccc 1610
|||||
Db 1530 CACGTTTCTTCCCC 1543
RESULT 4
D50405
LOCUS D50405 2091 bp mRNA linear PRI 10-FEB-1999
DEFINITION Human mRNA for RPD3 protein, complete cds.
ACCESSION D50405
VERSION D50405.1 GI:1665722
KEYWORDS RPD3 protein.
SOURCE Homo sapiens fetal lung cDNA to mRNA, clone:hRPD3.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 2091)
REFERENCE
AUTHORS Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (22-APR-1995) Yusuke Nakamura, Institute of Medical
Science, The University of Tokyo, Laboratory of Molecular Medicine,
Human Genome Center, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639,
Japan (E-mail:y-daigo@hms.u-tokyo.ac.jp, Tel:03-5449-5372,
Fax:03-5449-5433)
2 (bases 1 to 2091)
REFERENCE
AUTHORS Nakamura, Y.
JOURNAL Unpublished (1996).
REFERENCE
AUTHORS 3 (sites)
Furukawa, Y., Kawakami, T., Sudo, K., Inazawa, J., Matsumine, A.,
Akiyama, T. and Nakamura, Y.
TITLE Isolation and mapping of a human gene (RPD3L) that is homologous
JOURNAL to RPD3, a transcription factor in Saccharomyces cerevisiae
MEDLINE Cytogenet. Cell Genet. 73 (1-2), 130-133 (1996)
FEATURES
source
location/Qualifiers
1..2091
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p34.1"
/clone="hRPD3"
/tissue_type="fetal lung"
64..1512
/codon_start=1
/product="RPD3 protein"
/protein_id="BAA08909.1"
/db_xref="GI:1665723"
/translation="MAQTQGTTRKVCYYDDGVGNYYYGGHPMKPHRIRMTNLLLN
YGLYRKMEIYRPHKANAEMTKYHSDYIKFLSIRPDNMESEYSKOMRFNVGDCPV
FDGLFEFCOLSTGGSVAVKLTKQOTDIAYNMGGLHAKKSEASGFCYVNDIVLAI
LELKYHORVLYIDIDIHGDGVEAFYTTDRVYTFVSEFHKYGEYFPGTGLRDIGAGK
GKYAVNYPRLDGLDDESYEAIFFKPVMSKVMEMFQPSAVVLQCGSDLSGDRIGCFNL
TIKGAKCVEEFVSENLPLMLGGGTYTIRNVARCPTYETAVALDTEIPNELPYNDYF
EYFGPDFKLHISPMNTNONTNEYLEKIKQRLFENLRMLPHAPGVQMOAIPEDAIPEE
SGDEDEDPKRIRISICSSDRICACEEESDSEEGEGGRKNSNFKKAKRVTDEDEKE
KDPEEKVEVTEERTKEEKPEAKGVKEEVKLA"
polyA_signal 2073..2078
BASE COUNT 552 a 485 c 544 g 510 t
ORIGIN
Query Match 93.0%; Score 1498; DB 9; Length 2091;

Yeast

Best Local Similarity 99.3%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 97 gggagcgagcaagatgycgcagacgcagcggcaccggagagaaagtctgtactactag 156
|||||
Db 50 GGGAGCGGAGCAAGATGGCGCAGACGCGAGCGGACCCGGAGGAAGTCTGTACTACTAG 109
QY 157 acgggatgttggaaattactattatgacaagggcccaatgaagcctcaccgaatcc 216
|||||
Db 110 ACGGGGATGTTGGAATTACTATTATGACACAAGGCCACCCCAATGAGACCTCACCGAATCC 169
QY 217 gcataactataatttgcctcaactatgctcctaccgaaaaatggaatctatcgcc 276
|||||
Db 170 GCATGACTCATTAATTGCTGCTCAACTATGGTCTCTACCGAAAAATGGAATCTATCGCC. 229
QY 277 ctcaacaagccaatgctgagagatgaccgaagtlaccacagcgatgactacatlaattct 336
|||||
Db 230 CTCACAAAGCCAAATGCTGAGGAGATGACCAAGTACCACAGCGATGACTTAATTCT 289
QY 337 tgcgctccatccgtccagataacatgtcgaggtacagcaagcagatgcagagatcaacg 396
|||||
Db 290 TGGCCTCCATCCGCTCCAGATAACATGTCGAGTACAGCAAGCAGATTCACAGATTCAACG 349
QY 397 ttggtgagagactgtccagatattcgatgacctgttgagttctgtcagttgtactagt 456
|||||
Db 350 TTGGTGAGAGACTGTCCAGTATTCGATGGCCTGTTGAGTCTGTCACTGTCTACTGCTG 409
QY 457 gtctctgtgcaagtgtctgtgaacttaataagcagcagcagcgacatcgccgtgaattgg 516
|||||
Db 410 GTTCTGTGGCAAGTGTCTGTGAACCTTAATAAGCAGCAGACGAGACATCGCTGTGAATTGGG 469
QY 517 ctggggcctgcgaccatgcaagaagtlccgagcgatctgctctgttaacgtcaatgata 576
|||||
Db 470 CTGGGGGCGCTGCACCATGCAAGAAGTCCGAGGCATCTGGCTTCTGTACGTCAATGATA 529
QY 577 tcgtctgagccatcctggaactgctaaagatcaccaagaggtgctgtacattgacattg 636
|||||
Db 530 TCGTCTTGCCATCTCTGGAACCTGCTAAATATCACAGAGGGTGTGTACATTGACATTG 589
QY 637 atattcacatggtgaagcgctggaagggccttctacaccagcggatgacattgacattg 696
|||||
Db 590 ATATTACACCATGCTGACGCGCTGGAAGAGGCTTCTACACACGAGCGGCTCATGACTG 649
QY 697 tgccttcaataagatgagagacttcccaaggaactggggaactgaacgggataccggg 756
|||||
Db 650 TGTCTTTCATAAGTATGAGAGACTCTCCAGGAACCTGGGAGACCTACGGGATATCGGG 709
QY 757 ctggcaagaacaagtattatgctgttaactacccgctccgagacggatgtgacgagt 816
|||||
Db 710 CTGGCAAGGCAAGTATTATGCTGTACTACCCGCTCCGAGACGGGATGTGATGACGAGT 769
QY 817 cctatgagccatttcaagccggtcatgttccaaagtaatgagatgttccagcctagt 876
|||||
Db 770 CCTATGAGGCCATTTCAGCGCGGTCTATCTCAAGTAATGAGATGTTCCAGCCTAGTG 829
QY 877 cgtgtgtctacagttgtgctcagactccctatctgggatacgtttagtctgtcctcaatc 936
|||||
Db 830 CGGTGCTTACAGTGTGGCTCAGACTCCCTATCTGGGATCGGTAGGTTGCTTCAATC 889
QY 937 tatctataaagacacgcgaagtgtyggaattgtcaagagcttaacctgctatgc 996
|||||
Db 890 TAACTATCAAGGACACGCGCAAGTGTGTGGAATTGTCAAGAGCTTTAACTGCTATGC 949
QY 997 tgatgtctggagggcgtgtgttacaccaatctgtaacgttgccgggtgtgtgacatatgaga 1056
|||||
Db 950 TGATGCTGGAGGGCGGTGTACACCACTTCGTAACGTTGCCCGGTGCAGGACATATGAGA 1009
QY 1057 cagctgtgcccctgatacagagatccctaatgagcttccatatacactacttgaat 1116
|||||
Db 1010 CAGCTGTGGCCCTGGATACGAGATCCCTAATGAGCTTCCATATGACTACTTGAAT 1069
QY 1117 acttggaccagatttcaagctccacatcagctccttccaatatgactaaccagaacaga 1176
|||||

Db 1070 ACTTTGGACCAGATTTCACAGCTCCACATCAGTCCCTTCCAATATGACTAACCCAGAACAGCA 1129
QY 1177 atgaatccttgagaagaatcaaacagcgactgtttgagaaccttagaatgtctgcgcacg 1236
Db 1130 ATGAGTACCTGGAGAAGATCAACAGCGACTGTTGAGAACCTTAGAATGCTGCCGACG 1189
QY 1237 caacctggttccaacacgcagcgatctcctgagagacccatccctgagagagatgagatg 1296
Db 1190 CACCTGGGGTCCAAATGACGCGATTCCTGAGAGCGCATCCCTGAGAGAGAGTGGCGATG 1249
QY 1297 aggaacgaagacgacccctgacaagcgcatctcgatctgtcctctgacaacgaattgcct 1356
Db 1250 AGGACGAAGACGACCCCTGACAAGCGCATCTCGATCTGCTCCTTGACAAACGAATGGCT 1309
QY 1357 gtgaagaaagatctccgattctgaagagagagagggggccgcgaagaactctcca 1416
Db 1310 GTGAGGAAGAGTCTCCGATCTTGAAGAGAGAGGAGGGGGCCGCAAGACTCTTCCA 1369
QY 1417 acttcaaaaaagccaagagatcaaaaacagagatgaaaaaagaaaagaccagagagaga 1476
Db 1370 ACTTCAAAAAAGCCCAAGAGAGTCAAAAACAGAGATGAAAAAGAAAGACCCAGAGAGA 1429
QY 1477 agaaaaggaatcaccgaaagagagaaaaaccaaaggaagaaagccagaagccaaagggtca 1536
Db 1430 AGAAAGAGTCAACCCGAGAGAGAGAAAAACCAAGAGAGAAAGCCAGAAAGCCAAAGGGTCA 1489
QY 1537 aggaagagagccaagtgtgcctgaatgagacctctccagctctgtcctgtcgtgagtcct 1596
Db 1490 AGGAGGAGGTCAAGTGGCTGAATGAGACCTCTCCAGCTCTGGCTTCTGCTGAGTCCCT 1549
QY 1597 caagttctcttccc 1610
Db 1550 CACGTTTCTTCCCC 1563

RESULT 5
LOCUS AR012007 2111 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 2 from patent US 5763182.
ACCESSION AR012007
VERSION AR012007.1 GI:3969997
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2111)
AUTHORS Nakamura,Y. and Furukawa,Y.
TITLE RPD1 protein and DNA encoding the same
JOURNAL Patent: US 5763182-A 2 09-JUN-1998;
FEATURES
source 1. 2111
BASE COUNT 572 a 485 c 544 g 510 t
ORIGIN

Query Match 93.0% Score 1498; DB 6; Length 2111;
Best Local Similarity 99.3% Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 97 gggagcggaagaatgagcgcagacgcagggcaaccggaaggaagtctgttactactacg 156
Db 50 GGGAGGCGAGCAAGATGGCGCAGACGCGAGGGCACCCGGAGGAAGTCTGTACTACTACG 109
QY 157 acgggagatgttgaaattactatlatygacaagggccaaccaatgaagcctcacgaatcc 216
Db 110 ACGGGATGTTGAAATTACTATTATGACAAGGCCAACCAATGAAGCCTCACCGAATCC 169
QY 217 gcatgactcataattgtctgctcaactatgtctctaccgaaaaatggaatctatcgcc 276
Db 170 GCATGACTCATATTTGCTGCTCAACTATGCTCTCTACCGAAAAATGAATCTATCGCC 229
QY 277 ctcaacaagccaatgtcgtgagagatgaccaagtaccacagcgatgactacataaattct 336

Db 230 CTCACAAAGCCCATGCTGAGGAGATGACCAGATACACAGCGATGACTACATTAATTTCT 289
QY 337 tgcgctccatccgtccagataatacatgtcggagtacagcaagcagatgcagagattcaacg 396
Db 290 TGCGCTCCATCCGTCCAGATAACATGTCGGAGTACAGCAAGCAGATGCAGAGATTCAACG 349
QY 397 ttggtgaggaactgtccagatttcgatgtgcctgtttgagttctgtcagttgttactgtgtg 456
Db 350 TTGGTGAGGACTGTCCAGTATTCGATGGCCCTGTTGAGTTCTGTCTAGTTGTTACTGCTG 409
QY 457 gtctgtgcaagtgtctgaaacttaataagcagcagaacgacatcgccgtgaattggg 516
Db 410 GTTCTGTGGCAAGTGTGTGAACCTTAATAAGCAGCAGACGACATCGCTGTGAATTTGGG 469
QY 517 ctggggccctgcaccatgcaaaagaagtcaggagcatctgtctgttacctcaatgata 576
Db 470 CTGGGGCCCTGCACCATGCAAAAGAGTCCGAGGCAATCTGGCTTCTGTTACGTCAATGATA 529
QY 577 tcgtcttgccatctcgtgaactgtctaaagtatcaccaaggggtgtgtatcatgacatg 636
Db 530 TCGTCTTGCCATCCTGGAAGTCTGCTAAAGTATCACAGAGGGTGTGTACATTTGACATTG 589
QY 637 atattcaccatgtgtgaagcgctgtgaagagcgcttctacaccacgcgagccggtcatgactg 696
Db 590 ATATTCAACCATGTTGTGACGGCGGTGGAAGAGGCGCTTCTACACCACGGAGCCGGGTATGACTG 649
QY 697 tgtccttcatagaatgagagagtagtactccacgaactggygacctaagagataaccgggg 756
Db 650 TGTCTTTTCAATAGTATGAGAGAGTACTTCCAGGAAGTGGGACCTACGGGATATCGGGG 709
QY 757 ctggcaaaagacaagtattatgtctgttaactaccgcctccgagaacgagatgatacgagt 816
Db 710 CTGGCAAAAGGCAAGTATTTATGCTGTTAACCTACCCGCTCCGAGACGGGATGATGACGAGT 769
QY 817 cctatgagccatttcaagccggtcatgttccaaagtaatgagatgttccagcctagt 876
Db 770 CCTATGAGGCCATTTCAAGCCGGTCTCATGTCCAAAGTAATGAGATGTTCCAGCCTTAGTG 829
QY 877 cgtgtgtcttacagtggtgtcagactccctatcctgtggggtcgtgttagtgttcaatc 936
Db 830 CGGTGCTCTTACAGTGTGGCTCAGACTCCCTATCTGGGGATCGGTTAGTGTCTCAATC 889
QY 937 tatctataaagacacgcgaagtgtgtgaattgtcaagagccttaacctgtcctatgc 996
Db 890 TAACATCAAAAGGACACGCCCAAGTGTGTGAATTTGTCAAGAGCCTTAACCTGCCATATGC 949
QY 997 tgatgtcgtgagggcgtgtgttacaccatcgttaacglttgccggtgtcgtgacatatgaga 1056
Db 950 TGATGCTGGAGGCGGTGTTACACCATTCGTAACGTTGCCCGGTGCAGGACATATGAGA 1009
QY 1057 cagctgtgcccctgatacgygagatcccctaatyagcttccatacaatgactacttgaat 1116
Db 1010 CAGCTGTGGCCCTGGATACGGAGATCCCTAATGAGCCTTCCATACAAATGACTACTTGAAT 1069
QY 1117 actltgaccagatttcaagctccacatcagtccttccaatatgactaaccagaacaaga 1176
Db 1070 ACTTTGAGCCAGATTTCAAAGCTCCACATCAGTCTTCCAATATGACTAACGAAACACGA 1129
QY 1177 atgagtaccttgagaagatcaaacagcgactgtttgagaaccttagaatgtgcgcacg 1236
Db 1130 ATGAGTACCTGGAGAGATCAAAACAGCGACTGTTGAGAACCTTAGAATGCTGCCGACG 1189
QY 1237 caacctggttccaacgcagcgatctcctgagagacgccatccctgagagagatgagatg 1296
Db 1190 CACCTGGGGTCCAAATGCAAGCGGATTCCTGAGGACGCCATCCCTGAGGAGAGTGGCGATG 1249
QY 1297 aggaagaaagacccctgacaagcgcatctcgatctgtcctctgacaaaacgaattgcct 1356
Db 1250 AGGACGAAGACGACCCCTGACAAGCGCATCTGATCTGCTCTGTGACAAAAGAAATTTGCCCT 1309
QY 1357 gtgaagaaagatctccgattctgaagagagagagggggccgcgaagaactcttcca 1416

Db 1310 GTGAGAGAGAGTCTCCGATTCTGAAGAGAGGAGGGGGCCCAAGACTCTTCCA 1369
QY 1417 acttcaaaaaagccaagagagtcacaaacagagatgtaaaaaagaaaagaccagagagaga 1476
Db 1370 ACTTCAAAAAAGCCCAAGAGAGTCAAAACAGAGGATGAAAAAGAGAAAGACCAGAGAGAGA 1429
QY 1477 agaaaggaatcaccggaagagagaaaaaccaaggaaggaagcaaaaggggtca 1536
Db 1430 AGAAGAGAGTCAACCGAAGAGAGAGAAACCAAGGAGAGAGCCAGAAAGGGGTCA 1489
QY 1537 aggaagagcccaagttggcctgaatggacctctccagctctgtcttctgtctagtcct 1596
Db 1490 AGGAGAGGTCAAGTTGGCCTGAATGGACCTCTCCAGCTCTGGCTTCTGCTGAGTCCCT 1549
QY 1597 cagcttcttccc 1610
Db 1550 CACGTTCTTCCCC 1563

RESULT 6
US 52388
DEFINITION Sequence 2 from patent US 5659016.
ACCESSION 162388
VERSION 162388.1 GI:2480336
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2111)
AUTHORS Nakamura, Y. and Furukawa, Y.
TITLE RPD protein and DNA encoding the same.
JOURNAL Patent: US 5659016-A 2 19-AUG-1997;
FEATURES
source location/Qualifiers
1..2111
BASE COUNT 572 a 485 c 544 g 510 t
ORIGIN

no evidence
no primer

Query Match 93.0%; Score 1498; DB 6; Length 2111;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 97 gggagggcgaagaagatgagcgcagagcagggcaccggaggaagctgttactactag 156
Db 50 GGGAGGCGAGCAAGATGGCGCAGACGCGGACCCGGAGGAAGTCTGTACTACTACG 109
QY 157 acgggaggttggaaattactatattgacaaaggccaacccaatgaagcctcacgaatcc 216
Db 110 ACGGGGATGTTGAAATTACTATTATGACAAGGCCACCCCAATGAAGCCTCACGGAATCC 169
QY 217 gcatgactataattgtgtctcaactatgtctctaccgaaaaatgaaatctatcgcc 276
Db 170 GCATGACTCATATTGCTGCTCACTATGCTCTTACCGAAAAATGAAAAATCTATCGCC 229
QY 277 ctcaacaagccaatgtctgagagatgaccaagtlaccacagcgatgactacataatct 336
Db 230 CTCACAAAGCCCAATGCTGAGGAGATGACCAAGTACCACAGCGATGACTATAATTCT 289
QY 337 tggcgtccatccgtccagataacatgttcggagtacagcaagcagatgcagagattcaacg 396
Db 290 TGCGCTCCATCCGCTCCAGATACATGTCGAGTACAGCAAGCAGATGACAGATTCAACG 349
QY 397 ttggtgagactgttcagatattcgatggcctgtttgagttctgtcagttgtctactgtg 456
Db 350 TTGCTGAGGACTGTCAGTATTCGATGCGCTGTTGAGTTCTGTCACTTGTCTACTGCTG 409
QY 457 gtctgttggcaagtgctgtgaaacttaataagcagagacgcgcgtgaattggg 516
Db 410 GTTCTGTGGCAAGTGCTGTGAACCTTAATAAGCAGACGACGACATCGCTGTGAATTGGG 469
QY 517 ctggggcctgcacacatgcacaagaagtcgagcatctgtgtctgttaagtcacatgata 576

Db 470 CTGGGGGCTGCACCATGCAAGAAGTCGAGCATCTGGCTTCTGTTACGTCATGATA 529
QY 577 tgccttggccatccctggaactgtctaaagtalccacagaggtgtgtacattgacattg 636
Db 530 TCGCTTGGCCATCCTTGAAGTGTGAAGTATCACAGAGAGGGTGTGATGACATG 589
QY 637 atattcaccatgtgtgacggtgtggaagggccttctacacacagcagccgggtcatgtg 696
Db 590 ATATTCACCATGTTGACGGCGGTGAAGAGCCCTTCTACACACCGAGCCGGGTCTACTG 649
QY 697 tgtccttcaataagtatgagagttacttccacaggaacttgggagactacgggataccggg 756
Db 650 TGCTCTTCAATAGTATGAGAGTACTTCCAGGAACCTGGGACCTACGGGATATCGGGG 709
QY 757 ctggcaagacaagtalatgtgttaactaccgcctccgagacggatgtatgacagat 816
Db 710 CTGGCAAGGCAAGTATTATGCTGTTAACTACCCGCTCCGAGACGGGATGATGACGAGT 769
QY 817 cctatgagccatttcaagccggtcatgtccaaagtaatgagatgttccagcctagt 876
Db 770 CCTATGAGGCCATTTTCAAGCCGGTCAATGTCCTCAAGTATGAGATGTTCCAGCCTAGTG 829
QY 877 cgggtgtcttaccagtgtgtgtcagactccctatctgggagtcggtatgtgttcaatc 936
Db 830 CGGTGCTTACAGTGTGGCTCAGACTCCCTATCTGGGATCGGTGAGGTGCTTCAATC 889
QY 937 tatctatcaagagacaagccaagtggtggaattgtcaagagctttaacctgtccatg 996
Db 890 TAACATCAAGGACACGCCAAGTGTGTGAATTTGTCAAGAGCTTTAACTGCTATGC 949
QY 997 tgatgctgggaagcggtgtgttaccacactcgttaacggttgcgggtgtgtgacatatg 1056
Db 950 TGATGCTGGGAGCGGTGTTACACCATTCGTAACTGTGCCCGGTGACAGACATATGAGA 1009
QY 1057 cagctgtgcccctgtagtacgagatccctaatagagcttccatataactacttgaat 1116
Db 1010 CAGCTGTGGCCTTGATACGAGATCCCTTAATGAGCTTCCATACATGACTACTTGAAT 1069
QY 1117 acttggaccagatttcaagctccacatcagcttccatataatgataaaccagaacaga 1176
Db 1070 ACTTTGACCGAGATTTCAAGCTCCACATCACTCTTCCAATATGATTAACAGAACACGA 1129
QY 1177 atgagtaacctggagaagatcaaacagcagctgttgaaccttgaatgtgtgcgcag 1236
Db 1130 ATGAGTACCTGGAGAAGATCAAAACAGCGACTGTTGAGAACCTTAGAATGCTGCCGACG 1189
QY 1237 cactgtgggtccaaacgcagcgatctcctgtaggaagccatccctgtaggaagtgatg 1296
Db 1190 CACCTGGGTCCTCAATGACAGCGCATTCCTGAGGACGCCATCCCTGAGAGAGTGGCGATG 1249
QY 1297 aggaaggaagacccttgacaagcgcatctgactgtctcctctgacaaacgaattgct 1356
Db 1250 AGGACGAGACGACCCCTGACAAAGCCCATCTGATCTGCTCCTGCAACAAAGAAATTGCC 1309
QY 1357 gtgaaggaagttctccgattctgaagaggaaggaagggggccgcaagaactcttcca 1416
Db 1310 GTGAGGAAGAGTTCTCCGATTCTGAAGAGGAGGAGAGGGGGCCCAAGAACTCTTCCA 1369
QY 1417 acttcaaaaagccaagagagtcacaaacagagatgaaaaagagaagaaaccagagagaga 1476
Db 1370 ACTTCAAAAAAGCCCAAGAGAGTCAAAACAGAGATGAAAAAGAGAAAGCCAGAGAGAGA 1429
QY 1477 agaaaggaatcaccggaagaggaagaaaaaccaaggaaggaagccaaggggtca 1536
Db 1430 AGAAGAGAGTCAACCGAAGAGAGAGAAAAACCAAGAGAGAGCCAGAAAGGGGTCA 1489
QY 1537 aggaagagccaagttggcctgaatgagaccttccagctctgtcttctgtgagtcct 1596
Db 1490 AGGAGAGGTCAAGTTGGCCTGAATGAGACTCTCCAGCTCTGGCTTCTGCTGAGTCCCT 1549
QY 1597 cagcttcttccc 1610

Db	1550	CACGTTCTTCCCC	1563
RESULT	7		
EI1455			
ID	EI1455	standard; RNA;	HUM; 2111 BP.
AC	EI1455;		
XX			
SV	EI1455.1		
XX			
DT	07-OCT-1997 (Rel. 52, Created)		
DT	02-SEP-2000 (Rel. 65, Last updated, Version 2)		
DE	cDNA encoding human transcriptional control protein,RPDL.		
XX			
KM	JP 1996140687-A/1.		
XX			
OS	Homo sapiens (human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;		
OC	Eutheria; Primates; Catarrhini; Homidae; Homo.		
XX			
RN	[1]		
RP	1-2111		
RA	Nakamura Y., Furukawa Y.;		
RT	"RPDL PROTEIN AND DNA ENCODING FOR THE SAME";		
RL	Patent number JP1996140687-A/1, 04-JUN-1996.		
XX	JAPAN FOUND CANCER RES, EISAI CO LTD,		
OS	Homo sapiens (human)		
PN	JP 1996140687-A/1		
PD	04-JUN-1996		
PF	20-JUL-1995 JP 1995183763		
PR	22-SEP-1994 JP 94P 1227876		
PI	NAKAMURA YUSUKE, FURUKAWA YOICHI		
PC	C12N15/09,C07K14/39,C07K16/18,C12P21/02,C12P21/08,C12Q1/68,		
PC	(C12P21/02,		
PC	C12R1:19);		
CC	strandedness: Double;		
CC	topology: Linear;		
CC	hypothetical: No;		
CC	anti-sense: No;		
FH	key	Location/Qualifiers	
CC	FH		
CC	FT	source	1..2111
CC	CC	/organism="Homo sapiens"	
CC	FT	/tissue_type="Lung"	
CC	FT	/clone="L1-3977"	
CC	FT	1..63	
CC	FT	64..1512	
CC	FT	/product="human transcriptional control	
CC	FT	protein,RPDL"	
CC	FT	1513..2111	
XX	Key	Location/Qualifiers	
FH			
FT	source	1..2111	
FT		/db_xref="taxon:9606"	
FT		/organism="Homo sapiens"	
XX			
SQ	Sequence 2111 BP; 572 A; 485 C; 544 G; 510 T; 0 other;		
Query Match	93.0%	Score 1498; DB 23; Length 2111;	
Best Local Similarity	99.3%	Pred. No. 0;	
Matches 1504; Conservative	0; Mismatches 10; Indels 0; Gaps 0;		
QY	97	gggagcgagcaagatggcgcagacgcagcagccgcaggaaagtctgtactactacy	156
DB	50	GGGAGCGAGCAAGATGGCGCAGACGCAGGGCACCCGGAGGAAGTCGTACTACTACG	109
QY	157	acgggatgttggaattactatatgtgacaagcgaaccacatgaagcctcacgaatcc	216

D	b	110	ACGGGATGTTGAAATTA	TACTATTATGAGCAAGGCCACC	AATGAAGCCTCACCGAATCC	169		
O	y	217	gcattgacataa	tgtctgtccaactatggtc	ttaaccgaaaaatgaaatctatcgcc	276		
D	b	170	GCATGACTCAT	AATTGGTCTCAACTATGGTCT	CACGAAAAATGGAATCTATCGCC	229		
O	y	277	ctcaacaagcca	atgtctgaggagatgaaccaagt	taccacagcgatgactacatlaattct	336		
D	b	230	CTCACAAAGCCA	TGCTGAGAGATGACCAAGT	ACCACGCGATGACTACATTAATTCT	289		
O	y	337	tgcgctccatc	cgtccagataacatgttcg	agttacagcaagcagatgcagagattcaacy	396		
D	b	290	TGGCGCTCCAT	CCGTCAGATA	CATGTCGGAGTACAGACA	GATGCAGAGATTCAACG	349	
O	y	397	tttgttgaagga	ctgtccagtlatcgaatg	cgctgtttgagttctgtlcagtgtctactgtgtg	456		
D	b	350	TTGGTGAGGA	CTGTCCAGTATTGATGGCCTGTT	TGAGTTCTGTCAGTGTCTACTGGTG	409		
O	y	457	gttcctlygca	agtgctgtgaaaactlaataag	cagcagacgacatcgccgtgaattggg	516		
D	b	410	GTTCTGTGGCA	AGTGTGTGAACTTAATAAG	CAGCAGCAGACATCGCTGTGAATTGGG	469		
O	y	517	ctgggggcctg	caccatgccaagaagtlcc	gagcatctgtctgttacgtcaatgata	576		
D	b	470	CTGGGGGCCT	GCACCATGCAAGAAGTCCG	AGGCATCTGGCTTCTGTACGTAATGATA	529		
O	y	577	tcgtcttgcca	tccctgtgnaactgttaa	agtataccagaggytgctgtacattgacattg	636		
D	b	530	TCGTCCTTGG	CCATCTCGA	ACTGCTAAAGTATCACCA	GAGGGTGCTGTACATGACATTG	589	
O	y	637	aatttcaccaa	tgtgtgacgcygtggaagag	gccttctaaccaacygaccggttcattgactg	696		
D	b	590	ATATTCACCA	TGTTGATGACGGCGTGA	AGAGGCCCTTCTACACC	ACCGGCGGTCAATGACTG	649	
O	y	697	tgltccttcata	agtaatgaagagtlactl	cccaggaactyggacctacggyataccggyg	756		
D	b	650	TGTCCTTTCAT	AAGTATGAGAGTAGTACT	TCCCAGGA	CTGGGCACTACGGGATATCGGGG	709	
O	y	757	ctggcaaaaga	caaglatatgtctyltaacta	ccccgcctccgagacggygatgtgatbcyagt	816		
D	b	710	CTGGCAAAGG	CAAGTATTATGCTGTTAACT	ACC	CGCTCGAGACGGGATTGATGACGAGT	769	
O	y	817	cctaagagcca	atlttccaagccgblcata	gtlccaagaatgagatgttccaagcctagt	876		
D	b	770	CCTATGAGGC	CAATTTTCAAGCCGGTCAT	GTCCAAAAGTAA	TGAGATGTTCCAGCCCTAGTG	829	
O	y	877	csgtgtctcta	caqtgtgtgctcagactccc	atatctgsggatacgtttaggttgc	ttaac	936	
D	b	830	CGGTGCTT	TACAGTGTGGCTCAGACT	CCCTATCTGGGATCGGTTAG	GTGCTTCAATC	889	
O	y	937	tatctatcaaa	gagacacgcgccaaagt	tgttgyaatltgtcaagagc	ctttaacctgtcctatgc	996	
D	b	890	TAACTATFCA	AAGGACACGCCCAAGTGT	GGAA	TTGTCAAGAGCTTTAACCTGCCPATGC	949	
O	y	997	tgatgtctgga	ggcgcgtgtgttatcacca	atlcgttaacgttgc	ccggtgtctgacataatgaga	1056	
D	b	950	TGATGCTGG	GAGGCGGTGTTACACCA	TTTCGTAACGTTGCCCGGTG	CAAGACATATGAGA	1009	
O	y	1057	cagctgtggcc	ctgatagacyagatccc	ctaatagagcttccaatacaata	gactacttgyaat	1116	
D	b	1010	CAGCTGTGG	CCCTGGATACGAGATCC	CTAATGAGCTTCCATAC	ATGACTACTTTGAAT	1069	
O	y	1117	actttgagcca	gatttccaagctccacatc	agtlccttccaatalgactaac	cagaacaga	1176	
D	b	1070	ACTTTGAGAC	CAATTTTCAAGCTCCACAT	CAGTCCCTTCCAATATG	ACTAACCAGACAGACA	GA	1129
O	y	1177	atgagtacctg	gagaagatcaaaacg	cgaactgtttgagaa	ccttagaatgtcgcgcagc	1236	
D	b	1130	ATGAGTAC	CTGGAGAAGATCAAAAC	AGCGACTGTTGAGAA	ACCTTAGAATGCTGCCGACG	1189	
O	y	1237	caacttgggt	ctccaaacgacggcgat	ctcctgagagcgccatcc	ccttgagagagttgycgatg	1296	
D	b	1190	CACCTGGGG	TCCAAATGCAGCGGAT	TCTCCTGAGGACGCGCAT	TCCCTGAGGAGAGTGCGGATG	1249	

QY 1297 aggaacgaagacgaccctgacacgcgcatctcgatctgctcctctgacaaacgaatgcct 1356
|||||
Db 1250 AGGACGAAGACGACCCCTGACAGGCGCATCTCGATCTGCTCTTGACAAACGAATTGCCCT 1309
QY 1357 gtgaagaaagattctccgattctgaagagagagaggggcccgaagaactctcca 1416
|||||
Db 1310 GTGAGGAAGAGTCTCTCCGATTCTGACAGAGGAGGAGGGGGCCGCAAGAACTCTTCCA 1369
QY 1417 acttcaaaaaagccaagagagatcaaaaacagagatgtaaaaagagaagaccagagagaga 1476
|||||
Db 1370 ACTTCAAAAAAGCCCAAGAGAGTCAAAACAGAGGATGAAAAAGACAAGACCAGAGAGAGA 1429
QY 1477 agaaaggaatcacccgaagagagagaaaaacccaagagagagagaaagccaaggggtca 1536
|||||
Db 1430 AGAAGAAGTCAACCGAAGAGAGAGAAAAACCAAGGAGAGAGAACCAAGGGGTCA 1489
QY 1537 aggaaggaagccaagtgtgacctgaatggaacctctcagctctgtgcttctgtgagtcct 1596
|||||
Db 1490 AGGAGGAGGTCAAGTGTGGCTGAAATGGAACCTCTCCAGCTCTGGCTTCTGCTGAGTCCCT 1549
1597 caggttcttccc 1610
|||||
Db 1550 CACGTTCTTCCCC 1563

RESULT 8
LOCUS MMHIDE 1977 bp mRNA linear ROD 09-SEP-1997
DEFINITION M.musculus mRNA for histone deacetylase.
ACCESSION X98207
VERSION X98207.1 GI:1771285
KEYWORDS histone deacetylase.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1977)
AUTHORS Bartl,S., Taplick,J., Lager,G., Khier,H., Kuchler,K. and Seiser,C.
TITLE Identification of mouse histone deacetylase 1 as a growth
factor-inducible gene
JOURNAL Mol. Cell. Biol. 17 (9), 5033-5043 (1997)
MEDLINE 97415582
REFERENCE 2 (bases 1 to 1977)
AUTHORS Seiser,C.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-1996) C. Seiser, University of Vienna, Institute
of Molecular Biology, Vienna Biocenter, Dr.Bohr-Gasse 9, A-1030
Vienna, AUSTRIA

FEATURES
source
1.1977
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="Swiss 3T3"
18.1466
/codon_start=1
/product="histone deacetylase"
/protein_id="CAA66870.1"
/db_xref="GI:1771286"
/db_xref="MGI:108086"
/db_xref="SMISS-PROT:O09106"
/translation="MAQTQGTGRKVCYYVDGYNYYYGGHPMKPHRIRMTNLLN
YGLYRKMEIYRPHKANAEEMTKYHSDYIKFLRSIRPDNMEYSKOMQRFNVGDCPV
FDGLFEFCOLSTGGSVASAVKLNKOQTDIAVNWAGLHHAKKSEASGFCYNDIVLAI
LELKYHQRLVLYIDIDIHHGDVEEAFYTDVMTVSFHKYGEYFGTDLRDIGAGK
GKYAVVNYPLRDGIDDSYEAIKFPVMSKVMEMFQPSAVVLQCGSDLSGDRIGCFNL
TIKHAKEVEFVKSFNLPLMLGGGTYTIRNVAWCWTYETAVALDTEIPNDELPLYNDYF
EYFGPDFKLHISPSNMTNONTNEYLEKIKQRLFENLMLPHAPGVQMAIPEDAIPPE
SGDEDEEDPKRISICSSDKRIACEEESDSDEGEGRKNSNFKKAKRYKTEDEKE
KDPEEKVEVTEEEKTEKEKPEAKGVKEEVKLA"

BASE COUNT 543 a 475 c 488 g 471 t
ORIGIN

Query Match 77.0%; Score 1240.2; DB 10; Length 1977;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 1349; Conservative 0; Mismatches 163; Indels 1; Gaps 1;
QY 94 gcgggagcgagcaagaatgycgcagacgcagggcaccgcggaggaagtctgttactact 153
|||
Db 1 GCCGGGGGGGAGCAAGATGGCGCAGACTCAGGGCACCAAGAGAAAGTCTGTACTACT 60
QY 154 acgacgggagtggttggaaattactattatgtgacaaagcaccacccaatgaagcctcaccgaa 213
|||||
Db 61 ACGACGGGGATGTTGGAACACTACTATTATGGACAAGGGCACCCCATGAAGCCTCACCAGAA 120
QY 214 tccgcattgactataatttgcgtctcaactatgtgtctctaccgaaaaatggaatctatc 273
|||||
Db 121 TCCGCACTGACTCACAATTGCTGCTCAACTATGTGCTCTTACCAGAAATGGAGATCTACC 180
QY 274 gccctcaaaagccaatgctgtagagatgaaccaagtaccacagcagatgactacatlaaat 333
|||||
Db 181 GCCCTCACAAAGCCCAATGCTGAGGAGATGACCAAGTACCACAGTGAATGACTATTAAT 240
QY 334 tcttgctccatccgctccagataacatgctcgagtagacagacagatgcagagatcca 393
|||||
Db 241 TCCTGCGTTCTATTTCGCCAGATAACATGCTGTAATACAGCAAGCAGATGCAGAGATTCA 300
QY 394 acgttgtagagactgtccagatctcgatggcgtgttgagttcgtcgtcagttgtctactg 453
|||||
Db 301 ATGTTGGTGAGGACTGTCCGGTATTTGATGCTTGTGAGTCTGTCAGTTGTCCACGG 360
QY 454 gtgttctgtgcaagtgtgtgaacttaataagcagcagacgagacatgcgcgtgaatt 513
|||
Db 361 GAGGCTCTGTGCGCAAGTGTGTGAAGCTTAATAAGCAGCAGACGGACATCGCTGTAAC 420
QY 514 ggctgtaggagcctgcaccatgcacaaagaagtcgagacatcgtcttctgttactaatg 573
|||||
Db 421 GGGCTGGGGCTGCACCATGCAAGAGTCTGAAGCATCCGGCTTCTGTACGTCAATG 480
QY 574 atatgctcttgagcattcctggaactgttaagaatatacaccagagagtgctgtacattgaca 633
|||
Db 481 ACATCGCTTGGCCATCTCTGAAGTCTTAAGTACCACAGAGGGTCTCTATATTGACA 540
QY 634 ttgatattcaccatggtgacggtgtgaaagagggccttctacaccacgagccggtcatga 693
|||||
Db 541 TTGATATTCAACCATGGCGATGGCGTGGAAGAGGCCCTTCTATACTACAGACCGGTCATGA 600
QY 694 ctgtgtccttataagatatgagagtaactccacggaactgggagactaagcgatagcg 753
|||||
Db 601 CTGTGCTCTTCAATAATACGAGAGTACTTCCAGAACTGGGGACCTACGGGACATTGG 660
QY 754 ggctgtagcaagaacagatattatgctgttaactaccgcgtccgagacgagattgatgacg 813
|||||
Db 661 GGGCTGGCAAGGCAAGTACTATGCTGTGAACCTACCCACTGCGAGACGGCATTGACGACG 720
QY 814 agtccataagggcatttccaagccggtcatgtgccaaagtaatgagatgttccagccta 873
|||||
Db 721 AATCTATGAAGCCATCTTTAAGCCAGTATGTCCAAAGTAATGAGATGTTCCAGCCCTA 780
QY 874 gtgcggtgctcttaacagtgtgctcagactccctatctgagagatcggttagtcttca 933
|||||
Db 781 GTGACGTGTCTTACAGTGTGGCTCAGATTCCCTGTCTGGGAGCGGTTAGGTTGCTTCA 840
QY 934 atctatctatacaaaagacacgccaagtgtgtgaattgttcaagagctttaacctgccta 993
|||||
Db 841 ATCTGACCATCAAGAAGACACGCCCAAGTGTGGAGTTCGTGAAGATTTCACACTTGCCCA 900
QY 994 tgcgtatgcttgagagcggtgtttaaccattcgtaacgttgcccggtgtcgtgacatatg 1053
|||||
Db 901 TGCTGATGCTGGGAGGAGGTGGCTACACCATCCGGAATGTTGCTGCTGCTGACTTACG 960
QY 1054 agacagctgtggccctgatacagagatccctaatgagcttccatacaatgactacttg 1113
|||||
Db 961 AACACGCGGTGGCCCTGGACACAGAGATCCCTAATGAGCTGCCCTACATGACTACTTTG 1020

QY 1114 aatcttggaccagatttcaagctccacatcagctctccaatatagactaaccagaaca 1173
|||||
Db 1021 AATACTTTGACCGGATTTCAAGCTTCACATCAGCCCTTCTAACATGACCAACAGAACA 1080
QY 1174 cgaatgagctctgagaagaatcaacacagcagctgtttgagaacctagatgtccgc 1233
|||||
Db 1081 CTAACGAGTACTCTGAGAAGATCAAGCAGCGCTCTTTGAGAACTTGAGATGCTGCCCC 1140
QY 1234 acgcacctgggtcccaacgcagcgatctcctgagagccatccctgagagagtgcg 1293
|||||
Db 1141 ATGCCCTGGGTCAGATGAGGCCATCCCTGAGGAGCGCCATCCCGAAGAGAGTGGGG 1200
QY 1294 atgagcagcagcagccctgacacagcgcattctgattctcctctgacaaacgaattg 1353
|||||
Db 1201 ATGAGATGAGGAGGAGCCTTGACAAACCGCATCTCCATCTGCTCCTGTGATTAACGCATTG 1260
QY 1354 cctgtgaggaagattctcgcattctgaaagagagagagagggccgcaagaactctt 1413
|||||
Db 1261 CCTGTGAGGAGAGTCTCGGACTCAGATGAGGAGGAGAGGTGTCGCAAGAACTCTT 1320
QY 1414 ccaactcaaaaaagccaagagagatcaaaaagagatgaaaagagaagaaccagag 1473
|||||
Db 1321 CTAACCTCAAAAAAGCCAAAAGAGTTAAACAGAGGATGAGAAAGAGAAAGATCCTGAAG 1380
QY 1474 agaagaaagcaccagcagcagagagagaagaacccaagagagagagcagcagcagag 1533
|||||
Db 1381 AGAAAAAGAGATCACAGAAAGAGAGAAACCAAGGAGAGAGCCAGAAAGCCAAAGGG 1440
QY 1534 tcaagagagagcgaagtgtgcctgaatgagcctctccagctctgtgctgtctgagtc 1593
|||||
Db 1441 TCAAGAGAGAGGTCAAGTTGGCCTG-AGCAAGTCTGACAGCCCATCTTCCCAAGTT 1499
QY 1594 cctcacgtttctt 1606
|||||
Db 1500 CCTCACTCTCCCT 1512

RESULT 9
LOCUS MMU80780 1999 bp mRNA linear ROD 02-SEP-1997
DEFINITION Mus musculus putative histone deacetylase (HD1) mRNA, partial cds.
ACCESSION U80780
VERSION U80780.1 GI:2347179
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1999)
AUTHORS Johnson, C.A.
TITLE Putative histone deacetylase HD1 from M. musculus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1999)
AUTHORS Johnson, C.A., White, D., O'Neill, L.P. and Turner, B.M.
TITLE Direct Submission
JOURNAL Submitted (02-DEC-1996) Department of Anatomy, University of Birmingham, Birmingham B15 2TT, U.K.
COMMENT On Sep 2, 1997 this sequence version replaced gi:1737470.
FEATURES
location/Qualifiers
1..1999
/organism="Mus musculus"
/db_xref="taxon:10090"
1..1999
/gene="HD1"
15..1463
/gene="HD1"
/note="similar to human histone deacetylase HD1 and yeast RPD3"

CDS
gene
15..1463
/product="putative histone deacetylase"
/protein_id="AAB68398.1"
/db_xref="gi:2347180"
/translation="MAQTGQTKRKVCYYDDGVGNYYIGOGHPMXPHRIRMTNLLN

YGLYRKMEIYRPHKANAEMTKYHSDDIYIKFLRSIRPDNMSEYSKOMORFNVEDCPV
FDGLFEFCQLSTGGSVASAVKLNKQOTDIAVNMAGLHAKKSEASGFCYVNDIYLA
LELLKYHORVLYIDIDIHGDEVEEAFYTTDRVMVSPHKYGEYFPGTDLRDIGAGK
GKYAVNYPRLRDGIDDESEYEAIFKPYMSKVMEMFQPSAVVLQCGSDLSGDRIGCFNL
TIKHAACVEFKSFNPLMLGGGTYTIRNVARCWYETVALDTEIPNELPYNDYF
EYFGPDEKLIHISPSNMTNONTNEYLEKIKORLFENLMLPHAPGVOMAIPEDAIPBE
SGDEDEEDPKRISICSSDKRIACEEESDSDEEGEGRRKNSNFKKAKRVKTEDEKE
KDPEEKVEVTEEKTKEKPEAKGVKEVKLA"

BASE COUNT 583 a 467 c 480 g 467 t 2 others
ORIGIN
Query Match 76.7%; Score 1236.4; DB 10; Length 1999;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 1344; Conservative 2; Mismatches 163; Indels 1; Gaps 1;
QY 97 gggagcagcagaatgycgcagacgcagcagcccggaagaaagttgttactactacg 156
Db 1 GGGCGGAGCAAGATGGCGCAGACTCAGGGCACCAGAGGAAGTGTGTTACTACTACG 60
QY 157 acgggatgttgaaattactatattgacaagccaccgaatgaagcctcaccgaatcc 216
Db 61 ACGGGATGTTGAAACTACTATATGAGCAAGGGCACCCCATGARCGCTCACCGAATCC 120
QY 217 gcatactcataatttgcctcactatgtctcaccgaatgaagaaatgaaatctatgc 276
Db 121 GCATGACTCACAATTTGCTGCTCAACTATGCTCTACCGAAAAATGAGATCTACCGTC 180
QY 277 ctcaacaagccaatgtctgagagatgacccaagtaaccaagcagatgactacatlaattc 336
Db 181 CTCACAAGCCAAATGCTGAGAGATGACCAAGTACCACAGTGATGACTAATTAATTCC 240
QY 337 tgcgtccatccgtccagataacatgtctgagtagtacgacaagcagatgacagatccaacg 396
Db 241 TGGCTCTATTCGCCCAAGATAACATGTCTGATATACAGCAAGCAGATGACAGATTCAATG 300
QY 397 ttgtgagagactgtccagatattcgaatgctgttgaattctgtcagttgtctactgtg 456
Db 301 TTGCTGAGGACTGTCCGGTATTGTGATGGCTTTGTGAGTTCTGTCTGCTCCACGGGAG 360
QY 457 gtctgtgcaagtgtctgtgaacttaataagcagcagacgacatgcgcgtgaattgg 516
Db 361 GCTCTGTCCGAAGTGTGTGAAGCTTAATTAAGCAGACGACGACATGCTGTGAAGTGGG 420
QY 517 ctggggcctgcaccatgcaagaagtcgaagcattgtgcttctgttcaatgata 576
Db 421 CTGGGGCCTGCACCATGCAAAAGATCTGAAGCATCCGGCTTCTGTTACGTCATGACA 480
QY 577 tcgtcttgccatcctggaactgtctaaagtatcaccagaggtgtctatcatgacatg 636
Db 481 TCGTCTTGCCATCTCGGAAGTCTAAAGTACACACAGAGGGTGTCTATATGACATTTG 540
QY 637 atattcacatgtgacgcgtgtgaaagagccttctacaccagcagcgggtcatgactg 696
Db 541 ATATTCACCATGGCGATGGCGTGAAGAGCGCTTCTATACTACAGACCGGTCATGACTG 600
QY 697 tgtccttcaatagatgtgagagtagtctcccaaggaactgagggaactgaaggaatccgg 756
Db 601 TGTCCTTCAATAATACGAGAGTACTTCCCAAGAACTGGGGAGCTTACGGGACATTGGGG 660
QY 757 ctggcaagaagatattatgtgttaactaccgcctccgagacgggaattgtagcagagt 816
Db 661 CTGGCAAGGCAAGTACTATGCTGTGAAGTACCCACTGCGAGACGGCATTTGACGAGCAAT 720
QY 817 cctatgagccatlttcaagccggtcatgttccaagaagtaatgagatgttccagcctagt 876
Db 721 CCTATGAAGCCATCTTTAAGCCAGTCAATGCCAAAGTATGAGATGTTCCAGCCTAGTG 780
QY 877 cgtgtgtcttaagatgtgtgctcagactccctatctgaggatcgtgttaggttgcatac 936
Db 781 CAGTGTCTTACAGTGTGGCTCAGATTCCTGCTGCGGAGACCGGTTAGGTTGCTCAATC 840
QY 937 tatctataaaggacacgcgaagtggttgaaattgtcaagagccttaacctgcctatgc 996

Db	841	TGACCATCAAGGACACGCCAAGTGTGTGAGTTCGTGAAGAGTTTCAACTTGCCCATGC	900
QY	997	tgatctggaggcggtgtgttacaccattcgtlaacgttgcgcgtgtcgtgacatatgaga	1056
Db	901	TGATGCTGGAGGAGGTGGCTACACATCCGGAATGTTGCTCGCTGACTGACGAA	960
QY	1057	cagctgtggccctgtgatacggagatccctaagcttcatacaatgactattgaat	1116
Db	961	CAGCGGTGGCCCTTGACACAGAGATCCCTAATGAGCTGCCCTACAAAGACTATTGAAT	1020
QY	1117	actttgaccagatttcaagctccaacatcagtccttccaatatgactaacagacaga	1176
Db	1021	ACTTTGGACCGGATTTCAAGCTTCACATCAGCCCTTCTAACAATGACCAACAGACACTA	1080
QY	1177	atgagtaccttgagagaagatcaaacgcgactgttttgagaaccttagaatgtcgcgcacg	1236
Db	1081	ACGAGTACCTGGAGAGATCAAGACGCGTCTCTTTGAGAACTTGAGGATGCTGCCCATG	1140
QY	1237	cacctggggtcccaaacgcgacgcgattccttgaggagcgcacccctgaggagagtggcgatg	1296
Db	1141	CCCCGTGGGCTCCAGATGACAGGCCATCCCTGAGGACGCCATCCCGAAGAGAGTGGGATG	1200
QY	1297	aggacgaagacgaccctgacacgaagcgcattcgcattcgtccttgacaaacgaattgcct	1356
Db	1201	AAGATGAGAGGAGACCCTGACAAACCATCTCCATCTGCTCCTTGATAAACGCATTGCCCT	1260
QY	1357	gtgaggaaagttcttcgcattctgaagaggaggaggaggggcgccaagaactctcca	1416
Db	1261	GTGAGGAGAGAGTTCTCGAGACTCAGATGAGGAGGAGAGAGGTGGTCGCAAGACTCTTCTA	1320
QY	1417	acttcaaaaaaaccaagagagtcataaacagagatgtaaaaaagaaagaaagccagagagaga	1476
Db	1321	ACTTCAAAAAAGCCAAAAGAGTTAAACAGAGGATGAGAAAGAGAAAGATCCTGAAGAGA	1380
QY	1477	agaaaggaatcacccgaagagagagaaacccaagggagagaaagcccaagggagtgca	1536
Db	1381	AAAAAGAGTCACAGAGAGAGAGAAACCAAGGAGGAGAGAGCCAGAACCCAAAGGGGTCA	1440
QY	1537	aggaggaggccaagttggcctgaatgacctctcagctcgtgcttcctcgtgagtcct	1596
Db	1441	AAGAAGAGTCAAGTTGGCTG-AGCAAGGCTTGACGCCCATCTCTCCCAAGTTCCCT	1499
QY	1597	cacgttctctt 1606	
Db	1500	CACCTCTCCCT 1509	
RESULT	10		
AP000450			
LOCUS		81975 bp	DNA
DEFINITION		Homo sapiens chromosome 11 clone XXP1-307 map 11q23, WORKING DRAFT	HTG 30-MAY-2000
ACCESSION		AP000450	
VERSION		AP000450.3	GI:8118763
KEYWORDS		HTG; HTGS_PHASE1; HTGS_DRAFT.	
SOURCE		Homo sapiens DNA, clone:XXP1-307.	
ORGANISM		Homo sapiens	
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

Submitted (13-SEP-1999) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@gscc.riken.go.jp,

URL: <http://hgp.gsc.riken.go.jp/>, Tel:81-42-778-9923, Fax:81-42-778-9924)
On May 31, 2000 this sequence version replaced gi:7630227.

----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN

Web site: <http://hgp.gsc.riken.go.jp/>
Contact: hattori@gscc.riken.go.jp

----- Project Information
Center project name: HumDrafl1
Center clone name: XXP1-307

----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator EF-amersham; 100% of reads

Assembly program: Phrap; version 0.990329
Consensus quality: 79702 bases at least Q40
Consensus quality: 80980 bases at least Q30

Consensus quality: 81428 bases at least Q20
Insert size: 81575; sum-of-contigs
Quality coverage: 5.78x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 35184 contig of 35184 bp in length
35285 51779 contig of 16495 bp in length
51880 63010 contig of 11131 bp in length
63111 73886 contig of 10776 bp in length
73987 81975 contig of 7989 bp in length

Sequence updated (18-Apr-2000)
Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 35184: contig of 35184 bp in length
35185 35284: gap of 100 bp
35285 51779: contig of 16495 bp in length
51780 51879: gap of 100 bp
51880 63010: contig of 11131 bp in length
63011 63110: gap of 100 bp
63111 73886: contig of 10776 bp in length
73887 73986: gap of 100 bp
73987 81975: contig of 7989 bp in length.

Location/Qualifiers
1. 81975
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q23"
/clone="XXP1-307"

1. 35184
/note="assembly_fragment"
35285. 51779
/note="assembly_fragment"
51880. 63010
/note="assembly_fragment"
63111. 73886
/note="assembly_fragment"
73987. 81975
/note="assembly_fragment"

BASE COUNT 24973 a 15476 c 15771 g 25355 t 400 others
ORIGIN

Query Match		68.98;	Score 1110.4;	DB 2;	Length 81975;
Best Local Similarity		84.18;	Pred. No. 1.4e-290;		
Matches 1300;		Conservative	1;	Mismatches 237;	Indels 8; Gaps 4;
QY	67	gcgrctctgtcactcgggtagaccacgcgggagcgagcaagatgycgcagacgcagg	126		
Db	57895	GAGATTGACTGACAGAGAGACAGAGGACAGCAAGATGAGCAAGATGGCGCAAGGCAGG	57954		
QY	127	gcaccgcgaggaagctgttactactacgacgggatgttgaaattactattatgac	186		
Db	57955	GCACCAGAGAGAAAGTGTACTACTACGATAGGGGTGTGGAATTACTATATGAC	58014		
QY	187	aaggccaccaatgaagcctcaccgaatccgcgatctcataatttgtctgctcaactatg	246		
Db	58015	AAGCCACCCCAAGAAATCCTCACTGATCTGCGTGACTCATATTGTGCTGCTCACTATG	58074		
QY	247	gtctctaccgaaaaatggaatctatcgccctcacaagccaatgctgagagatgacca	306		
Db	58075	GTTTCTACTGAAAAATGGAACCCATCACCCCTCACAAAGCCAATGCTGAGGAGATGACCA	58134		
QY	307	agtaccacagcgatgactaattaaattcttgcgtcccatccgctccagataacatgctcg	366		
Db	58135	AGTACCACAGTGTGACTTAAATCTTGTGCTCCATCTATCCAGATTACGTGCTG	58194		
QY	367	agtacagcaagcagatgcagagattcaacgcttgytgcagactgtccagatctcgatggcc	426		
Db	58195	AGCACAGCAAGCGGATGCAGAGATTCAACGTTAGTAGAGACTGTCCAGCATTTGATGGCC	58254		
QY	427	tglttgagttctgtcagttgtctactggtgtctgtgcaagtgtgtgaaacttaata	486		
Db	58255	TGTTTGACTTGTGCTAGTGTCTACTGTGGCTGTGTGGCAAGTGCTGTGAACCTTAATG	58314		
QY	487	agcagcagacgacatgcgcgtgaattggcgtgggcctgcagccatgcagaagaagtcgcg	546		
Db	58315	AGCAGCAGATGACACTGCTGTGAATTGGGCTGGAGGCTTGACCCATGCAAAAGAGTTCA	58374		
QY	547	aggcatctgctctgttcaactatgatctgtcttggccatccctggaactgtcctaagt	606		
Db	58375	AGGCATCTGGCTCTGTATGTTAATGATACAAATCTGGGCCACCTTGAACTGCTAAAGT	58434		
QY	607	atcaccagaggtgtgttacatttgacatttgatatcccatggtgcagcgctggaagagg	666		
Db	58435	ATCACCAAGAGGTGTGTATGTGACGTTGATATTCACCATGATGATGGCATGGAGAAGG	58494		
QY	667	cctctacaccacgcgaccggtcatgactgtgtccctttcataagtatggagagtaactcc	726		
Db	58495	CCTTCTATATTACAGGCTAGGCCAAGACTGTGTCTTTCATTAATATGAGAGTACTTCC	58554		
QY	727	caggaactgggacctaacygataccggggctgcagaagaacagataltatgctgttaact	786		
Db	58555	CAGGAGATGGGACTTATGGGTTATTGGGGCTGGCAAGGCAGATATTATGCTATTAAC	58614		
QY	787	accgcctccgagacgggattgtagcagagtcctatgagccatttcaagccggtcatgt	846		
Db	58615	ATCTGCTCCTAGATGGATGTGATGATGAGTCCCTPATGAGGCCATTTTCAAACTGTCATGT	58674		
QY	847	ccaagaatagtgagtgtccagcctaagtgcggtgtcttacagtgtgtgctcagactccc	906		
Db	58675	CCAAAGTAATAGAGATGTCTCTAGTGTGGATCTTACAGCATGGCTCAGATGCCCT	58734		
QY	907	tatctggygatcggttagttgtcttcaatctatcatacaagagacagcccaagtgtgtg	966		
Db	58735	TGCTAGGGGTCACTTAGATTGTCTTAACTGTATCATCAAAAGGGCATGCCAAGTGTGTGG	58794		
QY	967	aattgtcaagagcttaacctgcctatgctgatatgctggagcggtgtgttacacacattc	1026		
Db	58795	AATCTGTCAAGAGTGTCAACCTGCCTATGTATTGATGGCAGAAAGGTGTTACACCATTT	58854		
QY	1027	gtaacgttgcccgtyg--ctgacatatgacagcagctgtggccctgatacagagatccc	1084		
Db	58855	ATAATGTTGCTCAGTGCCTCGAAATAGAAGAGAGCTGTGGCTCTGATATGAGATCCC	58914		
QY	1085	taatgagcttcacataatgactactttgaatactttggaccagatttcaagctccacat	1144		

Db	58915	TAATGAGCCTTCATACATGATGATTACTTTGAATACTTTGGACCATATTTCAAGCCCTACAT	58974
QY	1145	cagtccttccaatatgactaaccagaacacggaatgtagtacctgagagaagtccaacagcg	1204
Db	58975	CAGTCCTTCCAATATGACTAACCAGAACACTAATGATGACTCTGGAGAGAATTAACAGT	59034
QY	1205	actglttgagaaccttagaatgtctgcgcgacgacacctgggtccaaacgagcgattcc	1264
Db	59035	GCTGTTGAGAACCTTAGTAATGACAGGCCCATCCAGTCCAAATGCAAGTGATT-C	59093
QY	1265	tgaagcgcacatccctgaagagagtggcgatgaggaacgaagacccctgacaagcgcat	1324
Db	59094	TGAGGATGCTGTCCCTCAAGAGGGTGGCGATGAGAAATGAAGAGACCCTGACAA	59153
QY	1325	ctcgatctgctccttgacaacgaattgctgtgaggaagagtlctcgatctgaaga	1384
Db	59154	CTCCATCCGTTCTTGACAGATGAATTACCTGTGAGGAAGAGTCTTCGACTGTGATGA	59213
QY	1385	ggaaggagaggggcccgaagaactcttccaactcacaanaagccaagagagtcanaac	1444
Db	59214	GGAGGAGAGCGTGGCTGCAAGATCTCTCCATCTTCAAAAAAGCCAAGAGGATTTAAAC	59273
QY	1445	agaagatgaaaaagaagaagaccacagagagagaagaagaatcacccgaagagagaanaac	1504
Db	59274	AG---GATGAAAAAGACAGACTCAGAGATGAAGAAATGTCAACAGAGAGGAGA-AAAAT	59328
QY	1505	caaggaagagaagcccaagaagcgaagggctcaaggaaggaagcgaagttgctgaatga	1564
Db	59329	GAAAGAGGAGAAAGCCAGAGACCAAAAGAGGTCAAGGAGCAAGTCAAGCTGGCCTGAGGGA	59388
QY	1565	cctctcagctctgcttctcgtgtagtccctcacglttcttccc	1610
Db	59389	CCTCCCTGCTGTGCTTCTCTGCTAGTCCCAATCTTCTCTCCC	59434
RESULT 11			
AC020549			
LOCUS			
DEFINITION			
AC020549			
VERSION			
AC020549.4			
KEYWORDS			
HTG.			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 146841)			
Sulston,J.E. and Waterston,R.			
Toward a complete human genome sequence			
JOURNAL			
MEDLINE			
99063792			
2 (bases 1 to 146841)			
Edwards,J., Drone,K., Laplant,Y. and Ahluwalia,R.			
The sequence of Homo sapiens BAC clone RP11-172C16			
JOURNAL			
REFERENCE			
3 (bases 1 to 146841)			
Waterston,R.H.			
Direct Submission			
JOURNAL			
REFERENCE			
4 (bases 1 to 146841)			
Waterston,R.H.			
Direct Submission			
JOURNAL			
REFERENCE			
5 (bases 1 to 146841)			
Waterston,R.			
Direct Submission			
JOURNAL			
REFERENCE			
Submitted (08-NOV-2000)			
Department of Genetics, Washington			
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA			

REFERENCE	6 (bases 1 to 146841)
AUTHORS	Waterston, R.
TITLE	Direct Submission
JOURNAL	Submitted (09-MAY-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT	On Aug 11, 2000 this sequence version replaced gi:7630875.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplens@watsn.wustl.edu
----- Summary Statistics -----
Center project name: H_NH0172C16

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>.

SOURCE INFORMATION:
The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.reagen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-720G3; the clone sequenced to the right is RP11-643E23. Actual start of this clone is at base position 1 of RP11-172C16; actual end is at base position 146841 of RP11-172C16.

The region from base position 130951 to 131438 represents the best possible assembly of an imperfect TA repeat. Restriction digest information from EcoRV, HindIII, and BamHI indicates that approximately 150 to 250 bases may be missing from the final sequence.

```
FEATURES
source
    Location/Qualifiers
    1. .146841
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11"
    /clone="RP11-172C16"
    /clone_1fb="RPc1-11"
    418. .575
    /rpt_family="MER1_type"
    774. .946
    /rpt_family="L1"
    1114. .1313
    /rpt_family="L1"
    1365. .1452
    /rpt_family="ERV"
    1425. .1510
    /rpt_family="ERV"
```

repeat_region	1541. .2262	/rpt_family="L1"	EST AW009077 (NID:g5857855) ws76g06.x1"
repeat_region	2269. .2957	/rpt_family="L1"	EST AW844097 (NID:g7938080)"
repeat_region	8426. .8655	/rpt_family="L1"	EST AW841802 (NID:g7935785)"
repeat_region	8656. .8957	/rpt_family="Alu"	
repeat_region	8958. .8996	/rpt_family="L1"	
repeat_region	9029. .9453	/rpt_family="L1"	
repeat_region	10059. .10457	/rpt_family="MALR"	
repeat_region	10468. .10610	/rpt_family="L1"	
repeat_region	10659. .10807	/rpt_family="L2"	
repeat_region	10843. .11005	/rpt_family="ERV1"	
repeat_region	12544. .12725	/rpt_family="L1"	
repeat_region	12830. .13126	/rpt_family="Alu"	
repeat_region	13433. .13674	/rpt_family="L1"	
repeat_region	14419. .14587	/rpt_family="MIR"	
repeat_region	15199. .15545	/rpt_family="L2"	
repeat_region	16012. .16077	/rpt_family="L2"	
misc_feature	17372. .17741	/note="similar to	
misc_feature	17555. .18180	/note="similar to	
misc_feature	17696. .18060	/note="similar to	
repeat_region	17999. .18157	/rpt_family="L2"	
repeat_region	18220. .18324	/rpt_family="L1"	
repeat_region	18325. .18401	/rpt_family="Alu"	
repeat_region	18408. .18538	/rpt_family="Alu"	
repeat_region	18539. .19181	/rpt_family="L1"	
repeat_region	19454. .19661	/rpt_family="L1"	
repeat_region	19705. .20140	/rpt_family="L1"	
misc_feature	20162. .20624	/note="similar to	
misc_feature	20371. .20690	/note="similar to	
repeat_region	22353. .22774	/rpt_family="MALR"	
repeat_region	23334. .23715	/rpt_family="ERV1"	
repeat_region	23738. .23802	/rpt_family="MALR"	
repeat_region	23803. .24380	/rpt_family="ERV1"	
repeat_region	24381. .24774	/rpt_family="MALR"	
repeat_region	24925. .25100	/rpt_family="L2"	
repeat_region	25169. .25250	/rpt_family="MER53"	
repeat_region	25524. .25566	/rpt_family="L2"	
repeat_region	25567. .25668		

```
repeat_region      /rpt_family="U13"
                    26564. .26944
repeat_region      /rpt_family="L2"
                    26947. .27061
repeat_region      /rpt_family="L2"
                    28123. .28317
repeat_region      /rpt_family="L2"
                    29799. .30188
repeat_region      /rpt_family="L2"
                    30312. .30484
repeat_region      /rpt_family="MERL_type"
                    38493. .38692
repeat_region      /rpt_family="Alu"
                    38712. .38863
repeat_region      /rpt_family="L1"
                    39064. .39272
repeat_region      /rpt_family="L1"
                    39323. .39868
repeat_region      /rpt_family="L1"
```

Query Match	68.9%	Score 1110.4;	DB 9;	Length 146841;
Best Local Similarity	84.1%	Pred. No. 1.4e-290;		
Matches 1300; Conservative	1;	Mismatches 237;	Indels 8;	Gaps 4;

QY	67	gcgcrtctgtcactcgcggggtgtaaccacgcgcggggaagcgcagcaatggtgcgacgcagc	126
Db	81647	GAGATTGACTGACAGAGAGAGACAGAGGGGACAGGAAGTGAGCAAGATGGCGCAAAAGCAGG	81706
QY	127	gcacccggaggaagtctgttactactacgcgcgggagtggttggaattactattatgac	186
Db	81707	GCACCAGGAGGAAGTCTGTACTACTACGATAGGGGTGTGGAAATTACTATTATGGAC	81766
QY	187	aagccaccaatgaagcctcaccgaalcgcgcatgactcataattgtctcactatg	246
Db	81767	AAGCCACCCCAAGGAATCCTCACTGAATCTGCGTGACTCATATTGTCTCAACTATG	81826
QY	247	gtctctaccgaaaaatgtaattctatcgccttcacaaagccaatgtcagagatgacca	306
Db	81827	GTTTCTACTGAAAAATGGAAACCCATCACCCCTCACAAGCCAAATGCTGAGGAGATGACCA	81886
QY	307	agtaccacagcgatgactacaattaaattcttgcgcctccatccgltccagataacatgtcg	366
Db	81887	AGTACCACAGTGATGACTACATTAAATTCTTGCTGCCATCTATCCAGATTACGTGTCTG	81946
QY	367	agtacagcaagcagatgcagagatccaacglttggtgagactgttccagtattcgatgcc	426
Db	81947	AGCACACGACAGCGGATGCACAGATTCAACGTTAGTGAGGACTGTCCAGCATTTTGATGGCC	82006
QY	427	tgttttagttctgtcagttgttctactctgtgttctgtgccaagtgttgaacttaata	486
Db	82007	TGTTTGACTTCTGTCACTTGTTCTACTGTGTGGCTCTGTGGCAAGTCTGTGAAACCTTAATG	82066
QY	487	agcagcagaagcagatcgcgcgtgaattggyctggygcctgcacatgcaagaagtccg	546
Db	82067	AGCAGCAGATGGACACTGCTGTGAATTGGGCTGGAGGCGTGCACCATGCAAAAGAAGTTCA	82126
QY	547	aggcatctggtctctgttaacgtcaatgatatacgctcttggccatccctgnaactgtaagt	606
Db	82127	AGGCATCTGGCTTCTGTATGTATGTTAATGATACAAATCTGGGCCACCTTGGAACTGCTAAAGT	82186
QY	607	atcacccagaggtgtgtatcatatgacatgtatataccaatgrytgaacgycgtgaaagag	666
Db	82187	ATCACCCAGAGGTTGTGTATGTTGACGTTGATATTCAACCATGATGATGGCATGGAAGAGG	82246
QY	667	ccttctacaccacgcgaccggtcatgactgtgtcccttccataagtatygagagtactcc	726
Db	82247	CCTTCTATATTACAGGCTAGGCCACGAGACTGTGTCTTTCATATAATATGAGAGAGTACTTCC	82306
QY	727	caggaactggygacctaacygatacccggygctgycgaaagacaagtattatgctgttaact	786
Db	82307	CAGGAGATGGGGACTTATGGGTTAATTGGGGCTGGCAAAAGGCAGATATTATGCTATTAACT	82366
QY	787	accgcgtccgagaacggaattgatacgcaggtcctatgagaccatttcaagccggtcatgt	846

Db	82367	ATCTGCTCCTAGATGGGATTTGATGATGAGTCCCTATGAGGCCATTTTCAAACTGGTCAATG	82426
QY	847	ccaaagtaatgagatgttccagcctagtgcggtgttcttacagltgtgtcagactccc	906
Db	82427	CCAAAGTAATAGAGATGTTCTGCTCTAGTGTGGTGATCTTTACAGCATGGCTCAGATGGCT	82486
QY	907	tatctgggagtcggttagttgtcttcaatctatctatacaagacacgccaagtgtgtg	966
Db	82487	TGTCTAGGGGTCACTTAGATTGCTTTAATCTGATCATCAAGGGGCATGCCAAGTGTGG	82546
QY	967	aattgtcaagagctttaacctgcctatgtctgatactggyagcggtgtgttacaccatc	1026
Db	82547	AATCTGTCAAGAGTGTCAACCTGCCCTATGTTATTTGATGGCAGAAAGGTGGTTACACCATTT	82606
QY	1027	gtaacgttgcgcggtg--ctggacatatgagacagctgtgyccttgatagcagagatccc	1084
Db	82607	ATAATGTTGCTCAGTGCCCTGGAAATAAGAAAGAGCTGTGGCTCTGATATGAGATCCC	82666
QY	1085	taatgagcttccatacaaatgactacttgaatactttggaccagatttcaagctccaat	1144
Db	82667	TAATGAGCCTTCATACAAATGATTACTTTGAATACTTTGACCATATTTCGAAGCCCFACAT	82726
QY	1145	cagtccttccataatgacttaaccagaacacgaatgagtaacctggagagaatcaaacgcy	1204
Db	82727	CAGTCCCTTCCAAATATGACTTAACCAAGAACACTAATGAGTACCTGGAGAAGATTAAACAGTG	82786
QY	1205	actgtttgagaaccttagaatgtctgcgcgacgcacacttgyggttccaaacycagcgatcc	1264
Db	82787	GCTGTTTGAGAACCTTAGTAATGACAGGCCCATCCACTGCAGTCCAAATGACAGGTGATT-C	82845
QY	1265	tgaggacgccaatccccttgaagagagtgycgatgagagcaagaacyacccttgacaagcycat	1324
Db	82846	TGAGGATGCTGTCCCTCAAGAGGGGTGGCGATGAGATGAAGAAGAACCCTGACAAGCTTCAT	82905
QY	1325	ctcgatctgtcctcttgacaacaacyaattgacctgtgagaaagagttctccgattctgaaga	1384
Db	82906	CTCCATCCGTTTCTCTGCACAGATGAATTACCTGTGAGAGAGGTTCTTGACTCTGATGA	82965
QY	1385	ggagggagaggggycgcgaagaactcttccaacttcaaaaaagccaagagagttcaaac	1444
Db	82966	GGAGGAGAGCGTGGCTGCACAGATCTCTCCCATCTTCAAAAAAGCCCAAGAGGATTAAAC	83025
QY	1445	agagatgaaaaagagaagaaccacagagagagaagaagaatcacccgaagagagaanaac	1504
Db	83026	AG----GATGAAAAAAGAGAGACTCAGAGATGAAGAATACTCACAGAGAGAGAGA-AAAAAT	83080
QY	1505	caagggagagaagcccaagaagccaagaaggggtcaaaaggagagggccaagtgtgacctgaatga	1564
Db	83081	GAAAGGAGAGAGACCACGAGACCCAAAGAGGTCACAGAGACAGTCAAGCTGGCCCTGAGGGAA	83140
QY	1565	cctctcagactctggtcttctctgtgagtcacctcaagtttcttccc	1610
Db	83141	CCTCCCTGCTCTGCTCTCTCTGCTGCTAGTCCCAATCTTTCTCTCC	83186

RESULT	12			
LOCUS	AC068041/c			
DEFINITION	AC068041	187852 bp	DNA	linear
ACCESSION	AC068041			
VERSION	AC068041.4	GI:9965027		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	1 (bases 1 to 187852)			
JOURNAL	Waterston,R.H.			
REFERENCE	Unpublished			
REFERENCE	2 (bases 1 to 187852)			

Db 63751 TGTCTAGGGGTCACTTAGATGCTTTAATCTGATCATCAAGGGCATGCCAGTGTGCG 63692

QY 967 aattgtcaagagcttaactgtcctatgtctgctgagcggtgtgttacaccattc 1026

Db 63691 AATCTGTCAAGAGTGTCAACDCTGCCCTATGTTATGTGATGCGACAAGGTGTTACACCATTT 63632

QY 1027 gtaacgtgtcccggtg--ctggacatatgagacagctgtgcccctgatacagagatccc 1084

Db 63631 ATAATGTTGCTCAGTGCCTCGGAATAAGAAAGAGCTGTGGCTCTGTGATATGAGATCCC 63572

QY 1085 taatgagcttcacataactacttgaataactttggaccagatttccaagctccacat 1144

Db 63571 TAATGAGCCTTCATACATGATTAATCTTTGAATACTTTGAGCCATATTTCAGGCCCTACAT 63512

QY 1145 cagtccctccaatatgactaaccagaacagatgagtaacctgagagaagatcaaacacgcg 1204

Db 63511 CAGTCCCTCCAATATGACTTAACCAAGACACTAATGAGTACCTGGAGAAGATTAACAGTG 63452

QY 1205 actgtttgagaaccttagaattgctgcccagcagcactggggtcccaacgcagcattcc 1264

Db 63451 GCTGTTTGAGAACCTAGTAATGCAAGGCCCATTCACCTGCAGTCCAATGCAAGGTGATT-C 63393

QY 1265 tgaagcagccatccctcgtgagagagtgctgctgagagcgaagacccctgacaaagcgcac 1324

Db 63392 TGAGGATGCTGTCCCTCAGAGAGGGGTGGCGATGAGAAATGAAGAAGACCCCTGACAGCTCAT 63333

QY 1325 ctgcattctcctcctcgaacaaacgaattgctgtgaggaagagttctccgattctgaaga 1384

Db 63332 CTCCATCCGTTTCTCTGACAGATGAATTACCTGTGAGGAAGAGTTCTTGCATCTGATGA 63273

QY 1385 ggaagggaagaggggccgaagaactctccaactcaaaaaagccaagagagtcacaaac 1444

Db 63272 GGAAGGAGAGCGGTGGCTCAAGATCTCTCCCATCTTCAAAAAAGCCAAGAGAGATTAAAAAC 63213

QY 1445 agagagatgaanaagaaagaccagagagagaagaagaatcaccgaagagagagaac 1504

Db 63212 AG---GATGAAAAAGAGAGACTCAGAGATGAAGAAATGTACACAGAAGAGAGAGA-AAAAT 63158

QY 1505 caaggaagagaagccgaagaaggggtcaagagagaggaagccaagtgtgctgtgaatgga 1564

Db 63157 GAAAGGAGAGAGCCAGGAGCACAAGAGGTCAAGCAGCAGCAGTCAAGCTGGCCTGAGGGAA 63098

QY 1565 cctctccagctctgctctcctgtgagtcctccaagtttcttccc 1610

Db 63097 CCTCCCTGCTGTGGCTCTCCCTAGTCCCAATCTTCTCTCCC 63052

RESULT 13

LOCUS AC105705/c

DEFINITION Rattus norvegicus chromosome 1ddm4 clone CH230-67A16, WORKING DRAFT

AC105705

AC105705.1 GI:18092928

HTG: HTGS_PHASE1; HTGS_DRAFT.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 252957)

AUTHORS Muzny,D.M., Adams,C., Adlo-Oduola,B., Ali-osman,F.R., Allen,C., Alshrooks,S.L., Amatungue,H.C., Are,J.R., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,

Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Huijck,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratochvic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenko,S., Oguh,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 252957)

Worley,K.C.

Direct Submission

Submitted (09-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GNP

Center clone name: CH230-67A16

Summary Statistics

Assembly program: Phrap; version 0.990329first call to findphraplist

Consensus quality: 208655 bases at least Q40

Consensus quality: 216460 bases at least Q30

Consensus quality: 222170 bases at least Q20

Estimated insert size: 217567; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 70 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1

8561: contig of 8561 bp in length

8562

8661: gap of unknown length

8662

17870: contig of 9209 bp in length

17871

17970: gap of unknown length

17971

28613: contig of 10643 bp in length

28614

28713: gap of unknown length

28714

36948: contig of 8235 bp in length

36949

37048: gap of unknown length

37049

44586: contig of 7538 bp in length

44587

44686: gap of unknown length

* 44687 49960: contig of 5274 bp in length
* 49961 50060: gap of unknown length
* 50061 55765: contig of 5705 bp in length
* 55766 55865: gap of unknown length
* 55866 60143: contig of 4278 bp in length
* 60144 60243: gap of unknown length
* 60244 66193: contig of 5950 bp in length
* 66194 66293: gap of unknown length
* 66294 69993: contig of 3700 bp in length
* 69994 70093: gap of unknown length
* 70094 74621: contig of 4528 bp in length
* 74622 74721: gap of unknown length
* 74722 80106: contig of 5385 bp in length
* 80107 80206: gap of unknown length
* 80207 85208: contig of 5002 bp in length
* 85209 85308: gap of unknown length
* 85309 90190: contig of 4882 bp in length
* 90191 90290: gap of unknown length
* 90291 95100: contig of 4810 bp in length
* 95101 95200: gap of unknown length
* 95201 100892: contig of 5692 bp in length
* 100893 100992: gap of unknown length
* 100993 106570: contig of 5578 bp in length
* 106571 106670: gap of unknown length
* 106671 111667: contig of 4997 bp in length
* 111668 111767: gap of unknown length
* 111768 116678: contig of 4911 bp in length
* 116779 116778: gap of unknown length
* 116779 120909: contig of 4131 bp in length
* 120910 121009: gap of unknown length
* 121010 125487: contig of 4478 bp in length
* 125488 125587: gap of unknown length
* 125588 130004: contig of 4417 bp in length
* 130005 130104: gap of unknown length
* 130105 133645: contig of 3541 bp in length
* 133646 133745: gap of unknown length
* 133746 137029: contig of 3284 bp in length
* 137030 137129: gap of unknown length
* 137130 140398: contig of 3269 bp in length
* 140399 140498: gap of unknown length
* 140499 144725: contig of 4227 bp in length
* 144726 144825: gap of unknown length
* 144826 148959: contig of 4134 bp in length
* 148960 149059: gap of unknown length
* 149060 153396: contig of 4337 bp in length
* 153397 153496: gap of unknown length
* 153497 157396: contig of 3800 bp in length
* 157397 157396: gap of unknown length
* 157397 161939: contig of 4543 bp in length
* 161940 162039: gap of unknown length
* 162040 164650: contig of 2611 bp in length
* 164651 164750: gap of unknown length
* 164751 168117: contig of 3367 bp in length
* 168118 168217: gap of unknown length
* 168218 171870: contig of 3653 bp in length
* 171871 171970: gap of unknown length
* 171971 175124: contig of 3154 bp in length
* 175125 175224: gap of unknown length
* 175225 178279: contig of 3055 bp in length
* 178280 178379: gap of unknown length
* 178380 182598: contig of 4219 bp in length
* 182599 182698: gap of unknown length
* 182699 186637: contig of 3939 bp in length
* 186638 186737: gap of unknown length
* 186738 189183: contig of 2446 bp in length
* 189184 189283: gap of unknown length
* 189284 191324: contig of 2041 bp in length
* 191325 191424: gap of unknown length
* 191425 193902: contig of 2478 bp in length
* 193903 194002: gap of unknown length
* 194003 196283: contig of 2281 bp in length
* 196284 196383: gap of unknown length
* 196384 199868: contig of 3485 bp in length

* 199869 199968: gap of unknown length
* 199969 204016: contig of 4048 bp in length
* 204017 204116: gap of unknown length
* 204117 206274: contig of 2158 bp in length
* 206275 206374: gap of unknown length
* 206375 208084: contig of 1710 bp in length
* 208085 208184: gap of unknown length
* 208185 209301: contig of 1117 bp in length
* 209302 209401: gap of unknown length
* 209402 212290: contig of 2889 bp in length
* 212291 212390: gap of unknown length
* 212391 214423: contig of 2033 bp in length
* 214424 214523: gap of unknown length
* 214524 216573: contig of 2050 bp in length
* 216574 216673: gap of unknown length
* 216674 218887: contig of 2214 bp in length
* 218888 218987: gap of unknown length
* 218988 220288: contig of 1301 bp in length
* 220289 220388: gap of unknown length
* 220389 222948: contig of 2560 bp in length
* 222949 223048: gap of unknown length
* 223049 226012: contig of 2964 bp in length
* 226013 226112: gap of unknown length
* 226113 227915: contig of 1803 bp in length
* 227916 228015: gap of unknown length
* 228016 229644: contig of 1629 bp in length
* 229645 229744: gap of unknown length
* 229745 231544: contig of 1800 bp in length
* 231545 231644: gap of unknown length
* 231645 233502: contig of 1858 bp in length

Query Match 65.4%; Score 1054; DB 2; Length 252957;
Best Local Similarity 84.5%; Pred. No. 3.2e-275;
Matches 1279; Conservative 0; Mismatches 225; Indels 10; Gaps 8;

QY 94 gcgggagcgcgcaagatgycgcagacgcggcaccgcggaggaagctgttactact 153
Db 68403 GCCGGATGCTGAGCAAGATGCTGCAGCCCTCAGGGGTGTCAAGAGGAATTCGTACTGCT 68344
QY 154 acgaacggagatgttgaataattactatlatgacaagcaccaccaatgaagcctcaccgaa 213
Db 68343 AGGACTGGGATGTCGGAATACTACTATTATGACAAAGGGCACCACCAATGAACTTACTGAA 68284
QY 214 tccgcatactcataattgtctgtcacaatactgtctctaccgaaatgaatcctatc 273
Db 68283 TCCGCATGACTCAAAATTTGCTGTCCAACCATGCTCTTATTG-AAATGGAATCTTACC 68225
QY 274 gccctcacaagccaatgtctgagagatgaccaagtaccacagcgtactacatnaat 333
Db 68224 ATCCTCACAAGCCACTGCTGAGGAGATGACCAAGTACCACAGTGATCACTTAAT 68165
QY 334 tcttgccctccatccgtccagataacatgtctcgaggtacagcaagcagatcagagattca 393
Db 68164 TCTTGCACTCTATTCCGCCACAGATAATATGTGTAATATAGCAAGCAGATGCAGATTCA 68105
QY 394 acgttggtgagactgtccaglatcgtatgagcctgtttgagttctgtcagttgtcactg 453
Db 68104 ACAGGGGTGAGGTCGTGTCGGGTATTGATGCTGTTGAGTTTGTGCGTTGTCCACGG 68045
QY 454 gtggtctgttgcaagtgtctgtaacttaataagcagcagcagcagcgcgtgaatt 513
Db 68044 GTGGCTCTGT-CCAAGTGTCTGTAACCTCAATAAGCAGACAGACAGACATCGCTGTAAC 67986
QY 514 gggctggggcctgcaccatgcacaagaagtccgaggcacatctgctctctgttacatg 573
Db 67985 GCACCTGGGT-CTTGCAACCAACGCAAGAGTGTGAAGCATCTGGCTTCTGTACGTCAATG 67927
QY 574 atatcgtcttgccatcctcgtgaactgttaaatatcaccagagaggtctgttacattgaca 633
Db 67926 ATATCGTCTTGCCATCTCTGGAACCTGCTGAGCTATTCACCAAGAGGGTGTGAATATTGACA 67867
QY 634 ttgatatccaccatggtgacgcgctggaagaggccttctacaccagcagccgggtcatga 693
Db 693 ttgatatccaccatggtgacgcgctggaagaggccttctacaccagcagccgggtcatga 693

Db	67866	TTTATATTC	CCCGTGGCTA	TGGTGTG	AAGAGG	CGCTTCT	TATAC	AACAGAC	AGCTGG	TCATGA	67807
QY	694	ctgtgtc	cttcataa	gtatgag	agtactt	ccccga	gaactg	gggac	ctacg	gataccg	753
Db	67806	CTGTG	CTCTTTT	TAATA	TGAGAG	TACTCT	CCAGAA	CTGGG	ACCTAC	GGAATAT	67747
QY	754	gggctg	gcaaga	gaatata	tatgtct	gttaact	accgct	ccgag	acgga	tgatgacg	813
Db	67746	AGCGT	GCAAGA	CACAGT	ACTATG	CCGTTA	ACTACC	ACTCG	AGACG	GCAATTA	67687
QY	814	agtccta	tgaggc	catcttc	aaagcc	ggtcat	gtccaa	agtaat	gagatg	ttccagcta	873
Db	67686	AGTCC	TATGA	AGCCAT	CTTCA	AGCCACT	CATGT	CCAA	GTATG	AGATGT	67627
QY	874	gtgcg	gtgtgt	cttacag	tgtgtg	ctcagac	ctccct	atctg	gggac	tcggttagt	933
Db	67626	GTGCAG	TGATCT	TAGAGT	GGGCTC	AGATTCC	CT-TCT	GAGGAT	CGGT	TAGATTG	67568
QY	934	atctata	ctcaag	gaacg	cccaag	tgtgtg	taaltg	tcaag	agctt	taacctg	993
Db	67567	ATCTGA	CTCTCA	AAAGG	GCATG	CCCAAG	TGTGT	GAGTTC	GTAAG	ATTCAAC	67508
QY	994	tgctgat	ctg	ggagc	ggtgtg	ttacac	catcg	taacg	ttgccc	ggtgtg	1053
Db	67507	TGCTA	TGTTGG	GAGAG	GTGGCT	ACACCC	TCCGTA	TGTCA	CTCGG	TGTGAG	67448
QY	1054	agacag	ctgtg	gccctg	atac	gagatc	cccta	atgag	cttc	ataatg	1113
Db	67447	AAACAG	CTGTGG	-CCAG	GCCACA	GAGATCC	CTAA	AGCTAC	CCCTACA	ATGACTT	67389
QY	1114	aatactt	g	gaccag	atctca	agctcc	acatc	agtc	cttc	caatatg	1173
Db	67388	AATACT	T	TGGAC	CAGATTT	CAAGCTT	CACAT	CAGCC	TTCC	AATATAG	67329
QY	1174	C-gaat	g	gtacct	g	ggaag	atca	aaacag	cga	ctgtttg	1232
Db	67328	CTTA	TGA	TACCT	AGTA	GAAGAT	CAAG	CAGT	GGCA	CTTTG	67269
QY	1233	caacg	ac	ctg	gggtcc	aaacg	cagc	gga	cttc	ctg	1292
Db	67268	CATG	CTCCT	AGGGT	CCAATG	CAGGCC	ATACC	AGAG	GGTGC	CATCC	67209
QY	1293	gatgag	ga	cgaag	acac	ccctg	aca	agcg	catc	tgatctg	1352
Db	67208	GAGGA	AGAT	TAAG	AGAC	CCCTG	ACAA	CACAT	CTCTG	CTCTG	67149
QY	1353	gcctgt	g	gagga	aggttc	ctcga	ttctg	aaag	gagga	gggag	1412
Db	67148	GCCTG	T	GAGGA	GAATCT	CTGAT	T	CAGAT	GAGGA	GAAGG	67089
QY	1413	tccaact	t	caaaaa	aaagc	aaagag	agag	agag	agag	agag	1472
Db	67088	TCTAA	CTTCA	AAAAAG	CCAAAG	AGTCA	AAAA	CAGAG	ATGAG	AAAAA	67029
QY	1473	gagaag	a	gaag	aatc	acacg	a	gaag	a	gaag	1532
Db	67028	G---	AAAA	AGAG	CCAC	AGAG	AGAG	AGAG	AGAG	AGAG	66972
QY	1533	gtcaag	g	agag	gagc	caagt	gtg	cctg	aatg	agac	1592
Db	66971	GTC	AAAA	AAAA	AGAG	AGTCA	ATTG	GCCTG	-GG	CAAGG	66913
QY	1593	ccctc	a	cg	tttct	1606					
Db	66912	CCCTC	A	TTTCC	CT	66899					

RESULT 14 AF044169 1646 bp mRNA linear VRT 13-NOV-2001
LOCUS AF044169
DEFINITION Gallus gallus erythrocyte histone deacetylase (HDAC1) mRNA,
complete cds.
ACCESSION AF044169

VERSION AF044169.1 GI:2829213
KEYWORDS chicken.
SOURCE Gallus gallus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1646)
AUTHORS Sun,J.M., Chen,H.Y., Moniwa,M., Samuel,S. and Davie,J.R.
TITLE Purification and characterization of chicken erythrocyte histone deacetylase 1
JOURNAL Biochemistry 38 (18), 5939-5947 (1999)
MEDLINE 99249807
PUBMED 10231548
REFERENCE 2 (bases 1 to 1646)
AUTHORS Sun,J.M., Chen,H.Y. and Davie,J.R.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-1998) Biochemistry, University of Manitoba, 770 Bannatyne Avenue, Winnipeg, MB R3E 0W3, Canada
FEATURES
source location/Qualifiers
1. 1646
/organism="Gallus gallus"
/db_xref="taxon:9031"
/cell_type="definitive erythrocyte"
/dev_stage="15 day embryo"
/note="also in adult liver"
1. 1646
/gene="HDAC1"
48. 1490
/note="HDAC1"
/note="associated with nuclear matrices of primitive and definitive erythrocytes"
/codon_start=1
/product="erythrocyte histone deacetylase"
/protein_id="AAC00504.1"
/db_xref="GI:2829214"
/translation="MALTGTGTRKVCYYDDGVGNYYGCGHPMKPHRIRMTNLLN YGLYRMEIYRPHKANAEMTKYHSDYIKFLRSIRPDNMSEYSKOMQRFNVGEDCV FDGLFEFCQLSAGGSVASAVKLKQOTDIAVNMAGGLHAKKSEASGFCYVNDIYLA ILELKHORVLYIDIDIHGDVEAEFTTDRVMVTSFHYGEYFGTGLRDIGAGK GKYYAVNPLRDGIDDESYEATFKPVISKVMEFQPSAVYLQCGSDLSGDRLCFNL TIKGHAKVVEFKVSFNLPMLMGGGYTIRNVARCWTEYAVALDTEIPNELPYNDYF EYFGPDKLHISPSMNTQNTNEYLEKIKORLFENLRMLDPAVQOMPIPEDAVQED SGDEEEDPEKRISRINSKDSDEEFSDESEDEGEGRNVANFKKAKRVKTEEKE EEEKRDEKEEKAKEEKAEPKGVETKST"

BASE COUNT 461 a 346 c 452 g 387 t
ORIGIN

Query Match 61.9%; Score 997.4; DB 5; Length 1646;
Best Local Similarity 80.2%; Pred. No. 5.2e-260;
Matches 1185; Conservative 0; Mismatches 286; Indels 6; Gaps 1;

QY	83	gggtag	accacg	cgggag	gcgag	aagatg	gcgcag	acgacg	gagc	accgga	ggaagt	142
Db	20	GGGTCT	GCGGCT	GTGGG	AGGCG	AGAAAT	GGCCTG	ACGCG	AGGGG	AGACCG	GAAGT	79
QY	143	ctgttac	tactac	gagc	gggagt	gtgga	aatata	tatatg	gaca	agggcca	cccaatgaa	202
Db	80	CTGCTA	CTACTA	CGACG	CGGTG	ATTGGA	AACTATT	ACTATG	GCACAG	GGGCA	CCCATGAA	139
QY	203	gcctcac	cgaa	tcgcgat	gactcata	atttgt	gtctca	actatg	gtctc	taccg	aaat	262
Db	140	GGCCCA	CAGGAT	CCGGAT	GACCC	ACACAA	CACTACT	GTGTA	CTACG	CGCTT	ACAGGA	199
QY	263	ggaatc	tatcg	ccctc	aca	aagcca	atgctg	agag	gatg	acca	agtacc	322
Db	200	GGAGAT	ATATCG	CCCTC	ACAA	GGCGGA	ACGCGG	AGAGAT	GACCA	AGTACC	ACAGATGA	259
QY	323	ctacat	taaat	cttgc	gtccat	ccgctc	gcagata	acatg	tcg	gagta	cagcaag	382
Db	260	CTACAT	CAAA	TTCCTG	AGATCC	ATCCG	CCGAC	AGACA	ATGTCT	GAGTAC	AGCAAG	319

QY 383 gcagagattcaacgttggtgagagactgtccagtatctcgatggtcgtgtttagttctgtca 442
Db 320 GCAAGATTTAACGTCGGGAGAGACTGCCCTGTGTTCGATGGGCTGTTGAGTCTGTCA 379
QY 443 gttgtctactgtgtgtctctgtggaagtgctgtgtaaaacttaataagcagcagcgacat 502
Db 380 GCTCTCTGCTGGAGGCTCCGTTGCCAGCGCTGTGAAGCTGAACAAGCACAGACAGATAT 439
QY 503 cgcctgtgaattggcgtggggcctgcaccatgcaagaagtcaggcatctgtctctg 562
Db 440 TCGTGTGAATTGGCAGAGAGGCTTCACCACGCTAAGAAGTCGGAGGCTTCTGGCTTCTG 499
QY 563 ttaacgtcaatgatalcgtctgtgccaatccctggaactgtctaagtalcaccaagagtgt 622
Db 500 TTATGTCAACGATATGTCTGCGCTACTTGTGAGCTCTTAAGTATCACCAGAGGGTCT 559
QY 623 gtacattgacattgatattccacatgtgtgacggtgtggaagggcctctacaccagga 682
Db 560 GTATATTGACATTGATATTCACCATGGAGATGTTGTGAGAGAAAGCCTTCTATACCACAGA 619
683 ccggtcatgactgtgtctcttcataaagtatgagaagtaacttcccaggaactgggacct 742
620 CCGTGTGATGACCGCTGTCTTCATTATAAGTATGAGAGTACTTCCCAGGAACAGGGGACCT 679
QY 743 acgggataccggggtcgtgcaaaagacaaglatattatgtcttaactaccgcctccgagacg 802
Db 680 GCGGGACATTGGTGCAGGGCAAAAGCAATACTATGTCTCAACTATCCCTCCGGGATGG 739
QY 803 gattgatgacgagtcctatgagccatttcaagccgtcatgtccaaagtaatgagat 862
Db 740 GATTGATGATGAGTCCCTACGAGCAATATTCAAGCCGGTGAATCTTAAAGTGATGAGAGAT 799
QY 863 gtccagcctagtgcgtgtgtcttacagtgtgctcagactccctatctgggagtcgtt 922
Db 800 ATTCCAGCCTTAGTGACAGTGTCTCTGCAGTGGCGGTCCGATTCCTGTCCGGGACAGGCT 859
QY 923 aggtgtctcaatctatctatcaaaagacacagccaaagtgtgtgaaattgtcaagagctc 982
Db 860 GGGTGTGTTTAATCTGACCACATCAAGGTCATGCCAAGTGTGTGAGATTGTCAAAAAGTTT 919
QY 983 taacctgcctatgtctgtatgctggtggaagcggtgtgttacaccatctgtaacgttgcctgtg 1042
Db 920 TAAATTGCTTATGTCTGATGCTGGAGAGAGGTGGCTATACGATCCGCAACGTGGCCAGATG 979
QY 1043 ctggacatatgagacagctgtgtcccttgatatacggaagatccctaattagcctccatacaa 1102
Db 980 CTGGACCTATGAGACTGCTGTGGCTTTGGACACTGAGATCCCAATGAGCTCCCATATAA 1039
QY 1103 tgactactttgaactcttgaccagaattlcaagctccacatcagtccttccaatatgac 1162
Db 1040 TGACTATTTTGAAGTACTTCGGACAGACTTTAAGCTGCACATCAGTCCCTCAAAACATGAC 1099
QY 1163 taaccagaacaagatgagtactgtgagaagatcaaacagcgactgtttgaaaccttag 1222
Db 1100 CAACCAAGAAATACCAATGAGTATCTCGAGAAGATCAAGCAGCGTCTCTTTGAGAATCTGCG 1159
QY 1223 aatgtgtccgcacgcacacctgggtgccaaacgcagcgcatctctgagagacgcacacctga 1282
Db 1160 CATGCTGCCCTCATGCCCCCTGGCGCTCCAGATGCAGCCAAATTCCTGAGGATGCTGTTCAGGA 1219
QY 1283 ggaagagtggcgatgaggaaggaagcaccctgacaaagcgcatctcgatctgtcctctga 1342
Db 1220 AGACAGTGGGGATGAAGAAGAGAAGATCCTGAGAAAGCGCATTTCAATCCGCAATTCTGA 1279
QY 1343 caaacgaatttgcctgtgaggaaggttctcgaattctgaagagaggaggggagcg 1402
Db 1280 TAAGAGAATATCTGTGATGAAGATTTCTCTGACTCTGAAGATGAAGGGGAAGGAGGCG 1339
QY 1403 caagaactcttccaacttcaaaaaagccaagagagtcataaagcagagatgaaaaagaga 1462
Db 1340 CAAAAATGTTGCCAACTTTAAGAGGCCAAAGCGTGTGAAAAAGAGAGAGAGGAAAAAGAGGA 1399
QY 1463 agaaccagagagagaagaagaaatcaaccgaagagagagaanaaccaagagagagagccaga 1522

Db 1400 -----GGAGGAGAAGAGGATGAGAAAGAGAGGAAAAAGCAAAAGAGAGAAAGCCGA 1453
QY 1523 agccaaaggggtcaagagagagcgaagtggcctga 1559
Db 1454 ACCCAAGGGGTGAAGGAAGAGACAAATCCACCTAA 1490
RESULT 15
AF039751 1678 bp mRNA linear VRT 18-AUG-1999
LOCUS
DEFINITION Gallus gallus histone deacetylase-1 mRNA, complete cds.
ACCESSION AF039751
VERSION AF039751.1 GI:2791683
KEYWORDS
SOURCE Chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1678)
AUTHORS Takami,Y., Kikuchi,H. and Nakayama,T.
TITLE Chicken histone deacetylase-2 controls the amount of the IgM
H-chain at the steps of both transcription of its gene and
alternative processing of its pre-mRNA in the DT40 cell line
JOURNAL J. Biol. Chem. 274 (34), 23977-23990 (1999)
MEDLINE 99377029
REFERENCE 2 (bases 1 to 1678)
AUTHORS Takami,Y.
TITLE Direct Submission
JOURNAL Submitted (23-DEC-1997) Biochemistry, Miyazaki Medical College,
Kiyotake Kihara 5200, Miyazaki 889-16, Japan
FEATURES
source location/Qualifiers
1.1678
/organism="Gallus gallus"
/db_xref="taxon:9031"
/cell_line="DT40"
58.1500
/note="HD-1"
/codon_start=1
/product="histone deacetylase-1"
/protein_id="AAB96923.1"
/db_xref="GI:2791684"
CDS
/translation="MALTQGTGRKVCYYVDGNYVYGGQHPMKPHRIRMTNLLN
YGLYRKMEIYRPHKANAEEMTKYHSDYIKFLKSTRPDMNSEYSKQMRNVGEDCPV
FDGLFEFCOLASGSVSAVKLNKOQTDIAVNMAGLHHAKESEASGFCYNDIVLAI
LELLKYHQRVLVIDIDIHGDVEEAFYTTDRVMTVSFHKYGEYFPGTGLRIDGAK
GKYAVNYPRLRDGIDSESYEALFKPVISKVMEFQPSAVVLQCGSDSLSGDRLCENL
TIKHAKEVEFKSFNLPMLLGGGYTIRNVAQWTYETAVALDTEIPNELPYNDYF
EYFGPDFKLHISPSNMTNONTNEYLEKIKORLEFENRMLPHAPGVQMPIPEDAVQED
SGDEEEDPEKRISIRNSDKRISDEEEDSEDEGGGRKNVANFKAKRVTEEKE
EEKKDEKEEKEAKEEKAEPKVKETKST"
BASE COUNT 462 a 357 c 466 g 393 t
ORIGIN
Query Match 61.9%; Score 997.4; DB 5; Length 1678;
Best Local Similarity 80.2%; Pred. No. 5.2e-260;
Matches 1185; Conservative 0; Mismatches 286; Indels 6; Gaps 1;
QY 83 ggtgaccacgcggyggagcgagcaagatggcgacagcagcgcaaccggaagaagt 142
Db 30 GGGTCTGGCGCGGTGGGAGGAGGAAATGGCGCTGACGCGAGGGGACCAAGCGGAAAGT 89
QY 143 ctgttactactaacgagcggtgtgtgaaattactatattagacaagggccaaccaatgaa 202
Db 90 CTGCTACTACTACGACGCGTGTGGAACCTATTACTATGGCCAGGGGCAACCCCATGAA 149
QY 203 gcctcacgaatccgcatgactcataattgtctgctcaactatgtctctaccgaaaat 262
Db 150 GCCCCACAGATCCGATGACCCACAACTTACTGCTGAACCTACGGCCTTTACAGGAAGAT 209
QY 263 ggaatctatcgccctcacaaagccaatgtgagagatgaccaagtlaccacagcgatga 322

Db 210 GGAGATATATCGCCCTCACAAAGCGGAGAGATGACCAAGTACCACAGATGATGA 269
QY 323 ctacattaaattcttggtccatccgtccagataacatgtcggagtacagcaagcaat 382
Db 270 CTACATCAATTCCTGAGATCCATCCGCCAGACACATGCTGAGTACAGCAAGCAGAT 329
QY 383 gcagagattcaacgttgtgaggaactgtccagatattcgatggcctgttgagttctgtca 442
Db 330 GCAAGATTTAACGTCGGGAGGAGACTGCCCTGTGTGATGGGCTGTTGAGTTCTGTCA 389
QY 443 gtgtctactgtgtgtctgtgtgcaagtgtgtgtaacttaataagcagcagagcaat 502
Db 390 GCTCTCTGCTGGAGGCTCCGTTGCCAGCGCTGTGAAGCTGAACAAAGCAGACAGATAT 449
QY 503 cggcgtgaattggcgtggggcctgacacatgcaaaagttccgagcatctggtctctg 562
Db 450 TGCTGTGAATTGGCAGGAGGCTTCACCCAGCCTAAGAGTCGAGGCTTCTGGCTTCTG 509
QY 563 ttacgtcaatgatatactgttggccatccctggaactgtctaaagatatcaccagaggtgt 622
Db 510 TTATGTCAACGATATGTCTGCTGCTATCTTGAGAGCTCTTAAGTATCACAGAGGCTCT 569
QY 623 gtacattgacattgatataccatcgtgtgacggcgtggaagagccctctacacacgga 682
Db 570 GTATATTGACATTGATATCAGCATGAGATGTTGTGAGGAAGCCTTCTATACCACAGA 629
QY 683 ccgggtcatgactgtgtcttccataagtatgagagtaacttccacgaactgaggacct 742
Db 630 CCGTGTGATGACCGGTGCTCTTCAATAAGTATGAGAGTACTTCCAGAGACAGGGGACT 689
QY 743 acgggataccgggctgtgcaaaagacaaagtattatgtcttaactaccgcctccgagacg 802
Db 690 GCGGGACATTGGTGCAGGCAAGGCAAAATACTATGCTGTCAACTATCCCCCTCCGGGAT 749
QY 803 gattgatgacgaagtcctatgagggcatttcaagccggtcatgtccaaagtaatgagat 862
Db 750 GATTGATGATGAGTCTTACGAGGCAATATTCAGCCGCTGATATCTAAAGTATGAGAGAC 809
QY 863 gtccagcctagtcggtgtcttacagtggtgtcagactccctatctggtgagtcgt 922
Db 810 ATTCCAGCCTAGTGAGTGTCTGCTGAGTGGGTCGATCTCTGTCGGGGACAGGCT 869
QY 923 aggtgtctcaatctatctacaagacagcaagtggtgaaattgtcaagacct 982
Db 870 GGGTGTGTTTAATCTGACCATGAAGGTCAATGCCAAGTGTGAGATTGTCAAAAGTTT 929
QY 983 taacctgcctatgtgtgtgtgagggcgtgttaccacattcgtaacgttgcccggtg 1042
Db 930 TAATTGGCCTATGCTGATGCTGGGAGGAGGTGGCTATACGATCCGCAACGCTGGCCAGAT 989
QY 1043 ctggacatatgagacagctgtggccctgatacagagatccctaatagagcttccatacaa 1102
Db 990 CTGGACCTATGAGACTGCTGTGGCTTTGGACACTGAGATCCCAATGAGCTCCCATATATA 1049
QY 1103 tgactacttgaataacttggaaacagatttcaagctccacatcaatccttccaatatgac 1162
Db 1050 TGACTATTTGAGTACTTCGAGACCAGACTTTAAGCTGCACATCAGTCCCTCAACATGAC 1109
QY 1163 taaccagaacagaaatgatactgagaaagatacaaacagcagactgtttgagaacctag 1222
Db 1110 CAACCGAATATCCAAATGATATCTCGAGAAGATCAAGCAGCGTCTCTTGAGAAATCTGCG 1169
QY 1223 aatgtgcgcgacgcaactgtggtccaaacgcagcgaattcctgagagcgcacatccctga 1282
Db 1170 CATGCTGCTCATGCCCCCTGGCGTCCAGATGACGCCAATTCCTGAGGATGCTGTTCAGGA 1229
QY 1283 ggaagatgagcagaaagacgaacccctgacaaagcgaatctcgatctgtcctctga 1342
Db 1230 AGACAGTGGGATGAAGAGAAGATCTTGAGAACGCCATTTCATCCGCAATCTCTGA 1289
QY 1343 caaacgaattgctgtgaggaagattctccgattctgaagagagagagggggcg 1402

Db 1290 TAAGAGATATCCTGTGATGAAGAAATTCTCTGACTCTGAAGATGAAGGGGAAGGAGCG 1349
QY 1403 caagaactcttccaaactcaaaaaagccaagaggtcaaaaacagagatgaaaaaagaa 1462
Db 1350 CAAAAATGTTGCCAACTTTAAGAAGGCCAAGCGTGTGAAAAACAGAGAGGAAAAAGAGA 1409
QY 1463 agaccagaggaagaagaagaaatcacccgaagagagaaacccaagaggaagaaagcaga 1522
Db 1410 -----GGAGAGAGAAGAGATGAGAAAGAAGAGAAAAAGCAAAAGAGAGAAAAAGCGGA 1463
QY 1523 agccaaagggtcaaggagagggccaagtgtgacctga 1559
Db 1464 ACCCAAGGGGTGAAGAGAAGACAAAAATCCACCTAA 1500

Search completed: June 8, 2002, 12:37:02
Job time: 10915 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 09:41:03 ; Search time 279.1 Seconds
(without alignments)
9910.244 Million cell updates/sec

Title: US-09-817-538-2
Perfect score: 1611
Sequence: 1 atgtctgggtctctgcgcg.....tccctcacgtttcttcccc 1611

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
al number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802: *
1: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1596.6	99.1	1611	22	AAC89554 Human histone deac
2	1499.6	93.1	2163	21	AAF16188 Human prostate can
3	1498	93.0	2111	17	AAI12940 RPD1, transcrip
4	1434.6	89.1	1449	18	AAI86371 CDNA encoding a hi
5	1378	85.5	1582	23	AAI81302 DNA encoding novel
6	818	50.8	1997	24	ABI99512 Mouse ischaemic co
7	805.2	50.0	1985	20	AAI32066 Human METH2 relate
8	805.2	50.0	1985	22	AAC90323 U31814 CDNA clone.
9	805.2	50.0	1985	22	AAC89555 Human histone deac

c	10	729	45.3	2145	23	ABL01897	Drosophila melanog
	11	640.6	39.8	4755	23	ABL01896	Drosophila melanog
	12	502.6	31.2	1943	20	AAI90840	Maize histone deac
	13	501.8	31.1	2019	20	AAI90839	Maize histone deac
	14	499.6	31.0	1990	21	AAI85259	Rice histone deace
	15	498.2	30.9	1807	22	AAI80350	Nucleotide sequenc
	16	480.2	29.8	1854	21	AAI89556	Human histone deac
	17	477	29.6	1805	22	AAI858260	Soybean histone de
	18	460.8	28.6	1576	20	AAI90841	Maize histone deac
	19	459	28.5	1826	20	AAI90837	Maize histone deac
	20	434	26.9	1662	23	ABL04121	Drosophila melanog
	21	433.4	26.9	1609	21	AAC46043	Arabidopsis thalia
	22	433.4	26.9	1611	21	AAC33022	Nucleotide sequenc
	23	431.8	26.8	1800	22	AAI80351	Human histone deac
	24	413	25.6	586	22	AAI825097	Human ovarian PCR-
	25	413	25.6	586	22	AAI83742	Human ovarian PCR-
	26	402	25.0	576	22	AAI825098	Human ovarian PCR-
	27	402	25.0	576	22	AAI83743	Human ovarian tumo
	28	392.4	24.4	600	21	AAI16397	Human colon cancer
	29	386.8	24.0	402	22	AAI55632	Human breast tumou
c	30	369	22.9	3793	23	ABL04120	Drosophila melanog
	31	329.2	20.4	1475	20	AAI90838	Maize histone deac
c	32	318.2	19.8	2009	22	AAI17396	Human bone marrow
c	33	286	17.8	286	23	AAI57741	CDNA #417 encoding
	34	284	17.6	3826	23	ABL01928	Drosophila melanog
	35	235.8	14.6	241	22	AAI87345	Human breast cance
	36	235.8	14.6	241	22	AAI17915	Human breast cance
	37	201.6	12.5	1654	21	AAI293329	Human histone deac
	38	201.6	12.5	2046	22	AAI59530	Human polynucleoti
	39	200	12.4	1367	22	AAI75178	Nucleotide sequenc
	40	200	12.4	1682	22	AAI51122	Nucleotide sequenc
	41	195	12.1	379	18	AAI86372	Partial CDNA encod
	42	169.4	10.5	375	18	AAI86373	Partial CDNA encod
c	43	158.4	9.8	511	22	AAI70719	Human cervical can
	44	151.8	9.4	227	18	AAI86374	Partial CDNA encod
c	45	137.4	8.5	470	22	AAI71919	Human cervical can

ALIGNMENTS

RESULT 1	AAC89554	standard; DNA; 1611 BP.
ID	AAC89554;	
XX		
AC	AAC89554;	
XX		
DT	08-MAR-2001	(first entry)
XX		
DE	Human histone deacetylase HDAC-1 coding sequence.	
XX		
KW	Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;	
KW	HDAC-D; cell cycle; tumorigenesis; cancer; inhibitor; antisense;	
KW	gene therapy; ds.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200071703-A2.	
XX		
PD	30-NOV-2000.	
XX		
PF	03-MAY-2000; 2000WO-IP01252.	
XX		
PR	03-MAY-1999; 9906-0132287.	
XX		
PA	(METH-) METHYLGENE INC.	
XX		
PI	MacLeod AR, Li Z, Besterman JM.	
XX		
DR	WPI; 2001-016407/02.	
DR	P-PSDB; AAB49954.	
XX		
PT	Antisense oligonucleotide that inhibits expression of a histone	

PT deacetylase, useful for treating and/or alleviating the symptoms of
PT neoplasia, or for inhibiting neoplastic cell growth in an animal -
XX
PS Disclosure; Page 50-51; 125pp; English.

The present invention provides inhibitors of histone deacetylase enzymes such as HDAC-1, HDAC-2, HDAC-3, HDAC-4, HDAC-5, HDAC-C and HDAC-D. These inhibitors may be antisense strands or they may be compounds identified by contacting the enzyme with the compound and measuring the resulting enzyme activity. These inhibitors are useful for treating cancers and for identifying which histone deacetylase is involved in a neoplasia.

Sequence 1611 BP; 428 A; 385 C; 440 G; 358 T; 0 other;

Query Match	99.18;	Score 1596.6;	DB 22;	Length 1611;
Best Local Similarity	99.48;	Pred. No. 0;		
Matches 1602; Conservative	0;	Mismatches	9;	Indels 0; Gaps 0;

1 atgtctggggtctctgcgccgctgtgtctgtctgtctcccaactcgttcattccctgagaacaca 60
 |||||
 1 atgtctggggtctctgcgccgctgtgtctgtctgtctcccaactcgttcattccctgagaacaca 60

QY 61 gcttagcgrctctgtcactcggggrtagaccacgcgggagcgagccaagatctggcgaga 120
||||||| |||||||
Db 61 qcctaaqcgtctctgtcactcggggrtagaccacgcgggagcgagccaagatctggcgaga 120

QY	121	cgcaggcaccgcgaggaagtctgttactactacgaacgggatgttggaattactatt	180
Db	121	cgcacgqcaccgcgaggaatcttgttactactacgaacgggatatttqgaattactatt	180

QY 181 atgacacagccacccaatgtgaagccctcaccgaatccgcgatgactcataatttgcgtctca 240
|||||
Db 181 atgacacagccacccaatgtgaagccctcaccgaatccgcgatgactcataatttgcgtctca 240

QY 241 actatggtctctaccgaaaatggaatctatcgccctcacaagaagccaatgctgtaggaga 300
|||||
Db 241 actatgattcttaccgaaaatggaatctatcccccacacaaagccaatgctgtaggaga 300
|||||

QY 301 tgaccaagtlaccacagcatgactcataaattcttgcgtccatccgtccagataaca 360
|||||
Dh 301 ttaaccaaataccacacccaatgaactaatttaattcttcacgtccatccatccagataaca 360

OY		361	tgtcggagtaacgaagcagaatgcagagattcaacglttgtagagactgtccagtattcg	420
Dh		361	 tcttcgaatacacccaacgaattgcgnagattcaacctttgatgaagactatccaatattcn	420

421 atggccctgtttagtctgtcagttgtctactgtgtgtctgtgcaagtcgtgtgaac 480
|||||
421 atggccctgtttagtctgtcagttgtctactgtgtgtctgtgcaagtcgtgtgaac 480
|||||

QY 481 ttataagcagcagacatgcgcgtgaattggctgggctgcaccatgcaaga 540
|||||
481 ttataagcagcagcagacatgcgcgtgaattggctgggctgcaccatgcaaga 540

541 agtcgagcatctgctctgtttagcgaatgatatactgctcttgccatccctggaactgc 600

QY 601 taagtatcaccagaggtgtgtacattgacattgatattcaccatgttgacgycgttg 660

661 aagagcctctacaccacgagccggtcatgactgtgtccttcataagatgagagat 720

QY 721 attccaggaactgggaacctacgggataccggggctggaagacaagtattatgctg 780
|||||

QY 781 ttactaccgcgtccgagacgggattgatgcagagctctatgagccatttcaagccgg 840

Db 781 ttaactaccgcctccgagacgagattatgacgagtcctatgagggccattttcaagccgg 840

QY 841 tcatgtccaaagtaatgagagatgttccagcctaagtgcggtgtcttacaagtgtgctcag 900
|||||
841 tcatgtccaaagtaatgagagatgttccagcctaagtgcggtgtcttacaagtgtgctcag 900

Qy 901 actccctatctggggatcgtgttagtgcctcaatctatcatcaagaagacagccaagt 960
|||||
Db 901 actccctatctggggatcgtgttagtgcctcaatctatcatcaagaagacagccaagt 960

Qy 961 gtgtggaattgtcaagagctttaactgcctatgctgtagctgaggcggtgtttaca 102
|||||
Db 961 gtgtggaattgtcaagagctttaactgcctatgctgtagctgaggcggtgtttaca 102

```
QY 1021 ccattcgtacggtgcccggtgtctgacatatgagacagcgtgtgcccctgatacggaga 108
|||||
Db 1021 ccattcgtacggtgcccggtgtctgacatatgagacagcgtgtgcccctgatacggaga 108
```

[illegible]

QY	1141	acatcagtccttccaatatagtactaacgaacacagatggttacctggagaagatcaaac	120
Db	1141	acatcagtccttccaatatagtactaacgaacacagatggttacctggagaagatcaaac	120

QY	1201	agcgactgttttgagaaccttagaatgtctgccgacgcacccggggtlccaacgcgagcgca	126
Db	1201	agcgactgttttgagaaccttagaatgtctgccgacgcacccggggtlccaatgcagcgca	126

QY	1261	ttcctgaggaagccatcccttgaggaagtggcgatgaggaagaagacgacccctgacaagc	132
Db	1261	ttcctgaagcagccatcccttgaggaagatgacatgaggaagaaagacgaccccttacaagc	132

QY 1321 gcatctcgatctgctccctctgacaaacgaattgcctgtgagggaagattctccgattctg 138
|||||
Dh 1321 acatctcgaatctactcctctgaacaaacgaattgccctgtgagaaagaattctccattctg 138
|||||

Dh 1381 aaagagagggaagggggygcgcaagaactctccaacttcaaaaaaagccaaagtagtca 144
|||
Dh 1381 aagagaaaagacaccccccccccccaacaactcttcgaacttcaaaaaaacccaagaadtca 144

QY 1441 aaacagagatgtaaaaagagaaagaccagagagagaagaagaatcaccgaatgaggaga 150
|||||
DB 1441 aaacagagatgtaaaaagagaaagaccagagagagaagaagaatcaccgaatgaggaga 150
|||||

QY 1501 aaaccaagagagagaagccagaaagccaaaggygtcaagagagagycacaaglttgycctgaa 156
|||||
Db 1501 aaacaaagcagcagaaacacagaaacacaaatcgaatcaacagcaaatcaaatcttacctgaa 156
|||||

QY 1561 tggacctctccagctctggtctcgtgtagtcacctcaagttctctccc 1611
|||||
1561 tggacctctccagctctggtctcgtgtagtcacctcaagttctctccc 1611
tggacctctccagctctggtctcgtgtagtcacctcaagttctctccc 1611

RESULT 2

ID	AAF16188	standard; cDNA; 2163 bp.
XX		
AC	AAF16188;	

DT	13-MAR-2001	(first entry)
XX		
DE	Human prostate cancer antigen nucleotide sequence SEQ ID NO:623.	

KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular
KW vulnery; gastrointestinal; nephrotropic; antinefctive; gynaecologi

NM gastroenteric, pulmonary, cardiovascular, prostatic, uterine, gastric, intestinal, wound; infectious disease; ss.
 XX
 OS Homo sapiens.

|||||
Db 1571 cacgttcttcccc 1584

RESULT 3

AAT12940
ID AAT12940 standard; cDNA to mRNA; 2111 BP.

AC AAT12940
XX 09-SEP-1996 (first entry)
DT

DE RPDL transcriptional control protein cDNA.

XX Transcription; expression; control; diagnosis; study; cancer;
KW mammary; gastric carcinoma; chromosome 1p34.1; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
FT CDS 64..1512
FT /tag= a
FT /product= RPDL

FT 5'UTR 1..63
FT /tag= b
FT 3'UTR 1513..2111
FT /tag= c

PN EP708112-A1.

PD 24-APR-1996.

PF 21-SEP-1995; 95SEP-0114884.

PR 22-SEP-1994; 94JP-0227876.

PA (CANC-) CANCER INST.
(EISA) EISA CO LTD.

PI Furukawa Y, Nakamura Y;

DR WPI; 1996-202003/21.

DR P-PSDB; AAR88919.

PT New human transcriptional control protein RPDL - used to develop
PT prods. for study and diagnosis involving the protein, partic. for
PT gene analysis

XX Claim 3; Page 10-12; 15pp; English.

CC AAT12940 encodes a transcriptional control protein, RPDL, derived from
CC a human foetal lung cDNA library. The gene encoding RPDL is localised
CC at 1p34.1 on the short arm of chromosome 1, this is a region where a
CC deletion is recognised in mammary and gastric carcinomas. The RPDL
CC gene is an important gene and has been found to be expressed in all
CC the following human tissues: heart, kidney, liver, lung, pancreas,
CC placenta, skeletal muscle, large intestine, peripheral leukocyte,
CC ovary, prostate, small intestine, spleen, testis and thymus gland;
CC but is not present in brain tissue. RPDL cDNA exhibits homology with
CC the yeast transcription factor RPD3 and is believed to have a similar
CC function. RPDL and DNA encoding it can be used in studying the
CC interactions of the protein with DNA and other proteins in the cell.

XX Sequence 2111 BP; 572 A; 485 C; 544 G; 510 T; 0 other;

Query Match 93.0%; Score 1498; DB 17; Length 2111;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 97 gggagcgagcaagatggcgagcagcgagcgagcgaggaagtctgttactactacg 156
Db 50 gggagcgagcaagatggcgagcagcgagcgagcgaggaagtctgttactactacg 109

QY 157 acgggagatgttggaattactattatgtgacaaggccaccatgaagcctcaccgatcc 216
Db 110 acgggagatgttggaattactattatgtgacaaggccaccatgaagcctcaccgatcc 169
QY 217 gcatgactcataatttctgtctcgaactatgtctcttaccgaaaaatggaatctatcgcc 276
Db 170 gcatgactcataatttctgtctcgaactatgtctcttaccgaaaaatggaatctatcgcc 229
QY 277 ctcaacaagccaatgtctgagagatgacccaagtaaccaagcgatgactcataattct 336
Db 230 ctcaacaagccaatgtctgagagatgacccaagtaaccaagcgatgactcataattct 289
QY 337 tgcgctccatccgtccagataacatgtcgaagtaaccaagcagatgcagagattcaacg 396
Db 290 tgcgctccatccgtccagataacatgtcgaagtaaccaagcagatgcagagattcaacg 349
QY 397 ttggtgagactgtccagtattcgaatgtgacctgttgaattctgtcagtgttactagtgtg 456
Db 350 ttggtgagactgtccagtattcgaatgtgacctgttgaattctgtcagtgttactagtgtg 409
QY 457 gtctgtggaagtgctgtgaaacttaataagcagcagacgacatcgccgtgaattggg 516
Db 410 gtctgtggaagtgctgtgaaacttaataagcagcagacgacatcgccgtgaattggg 469
QY 517 ctgggggacctgcaccatgcaagaagtcggagacatctgtctctgttcaatgata 576
Db 470 ctgggggacctgcaccatgcaagaagtcggagacatctgtctctgttcaatgata 529
QY 577 tcgtcttgccatccctggaactgtctaagatcaaccagaggtgtgtatcatgacatg 636
Db 530 tcgtcttgccatccctggaactgtctaagatcaaccagaggtgtgtatcatgacatg 589
QY 637 atattcaccatgtgacgcggttggaagagccttctacaccagcagcggttcatgactg 696
Db 590 atattcaccatgtgacgcggttggaagagccttctacaccagcagcggttcatgactg 649
QY 697 tgtccttcatagaatgtagagtagtctccaggaactggggacctaagcgatataccgggg 756
Db 650 tgtccttcatagaatgtagagtagtctccaggaactggggacctaagcgatataccgggg 709
QY 757 ctggcaagaagcagatattatgtctgttaactaccgcctccagacgggatgtatgacagat 816
Db 710 ctggcaagaagcagatattatgtctgttaactaccgcctccagacgggatgtatgacagat 769
QY 817 cctatgagccatttccaagccggtcatgtlccaagaatgtagatgttccaagcctagtgtg 876
Db 770 cctatgagccatttccaagccggtcatgtlccaagaatgtagatgttccaagcctagtgtg 829
QY 877 cgggtgtcttacagtgtgtgtcagactccctatctgtgggacgtgttagtcttcaatc 936
Db 830 cgggtgtcttacagtgtgtgtcagactccctatctgtgggacgtgttagtcttcaatc 889
QY 937 tatctataaagcagcagccaagtgtgtgaattgtlcaagagccttcaactgtcctatgc 996
Db 890 taactataaagcagcagccaagtgtgtgaattgtlcaagagccttcaactgtcctatgc 949
QY 997 tgatgtgagagcggtgtgttacaccatcgtaacgltgcccgtgtcggacatatagaga 1056
Db 950 tgatgtgagagcggtgtgttacaccatcgtaacgltgcccgtgtcggacatatagaga 1009
QY 1057 cagctgtgacctgtgatacggagatccctaagcttccatacatgactacttgaat 1116
Db 1010 cagctgtgacctgtgatacggagatccctaagcttccatacatgactacttgaat 1069
QY 1117 actttgaccagatttcaagctccacatcagtccttccaatatgtactaaccagaacagga 1176
Db 1070 actttgaccagatttcaagctccacatcagtccttccaatatgtactaaccagaacagga 1129
QY 1177 atgagtaacctggagaagatcaacaacgagactgtttgagaaccttagaatgtgtccgcacg 1236
Db 1130 atgagtaacctggagaagatcaacaacgagactgtttgagaaccttagaatgtgtccgcacg 1189

QY 1237 cacttgggtccaaacgcagcgattccttgagagcgcacccctgaggagagtggcgatg 1296
|||||
Db 1190 cacttgggtccaaatgatcagcgattccttgagagcgcacccctgaggagagtggcgatg 1249
QY 1297 aggaagaaagacgacccctgacaagcgatctcgatctgctctctgacaacgaattgctt 1356
|||||
Db 1250 aggaagaaagacgacccctgacaagcgatctcgatctgctctctgacaacgaattgctt 1309
QY 1357 gtgaggaagagttctcgcattctgaagaggagaggaggggcccgaagaactcttcca 1416
|||||
Db 1310 gtgaggaagagttctcgcattctgaagaggagaggaggggcccgaagaactcttcca 1369
QY 1417 acttcaaaaaagccaagagagtcataacagagatgataaagaagaaccagagagaga 1476
|||||
Db 1370 acttcaaaaaagccaagagagtcataacagagatgataaagaagaaccagagagaga 1429
QY 1477 agaaaggaatcaccgaagagagagaaacccaagagagagagagccagaagccaaagggtca 1536
|||||
Db 1430 agaaaggaatcaccgaagagagagaaacccaagagagagagagccagaagccaaagggtca 1489
QY 1537 aggaaggaagccaagttggccttgatgacctctccagctctgcttctgtgagtcctt 1596
|||||
Db 1490 aggaaggaagtcagttggccttgatgacctctccagctctgcttctgtgagtcctt 1549

QY 1597 cagcttcttctccc 1610
|||||
Db 1550 cagcttcttctccc 1563

RESULT 4
AAT86371
ID AAT86371 standard; cDNA; 1449 BP.

AAT86371:
16-APR-1998 (first entry)

CDNA encoding a histone deacetylase (HDx) protein, designated HD1.

HDx
Histone deacetylase gene; HDx; HD1; HDx polypeptide; deacetylation; H3;
H4; cell differentiation; chromatin structure; cell cycle progression;
proliferative disorder; fibroproliferative disorder;
degenerative disorder; autoimmune disease; HDx inhibitor; ss.

OS Homo sapiens.

Key Location/Qualifiers
1..0

/*tag= a

WO9735990-A2.

02-OCT-1997.

26-MAR-1997; 97WO-US05275.

26-MAR-1996; 96US-0624735.

(HARD) HARVARD COLLEGE.

Hassig CA, Jamison TF, Schreiber SL, Taunton J;

WPI; 1997-489651/45.
P-PSDB; AAW29324.

New isolated histone deacetylase polypeptide(s) and genes - used to
develop products for modulating the proliferation, survival or
differentiation of cells, e.g. for treating tumours.

Claim 27; Pages 112-114; 159pp; English.

The present cDNA sequence encodes a novel histone deacetylase (HDx)
polypeptide, designated HD1. The HDx polypeptides are capable of

modulating proliferation survival and differentiation of cells. The
proteins are able to alter chromatin structure by deacetylating histones
such as H3 or H4. They have the ability to modulate gene transcription. The
influencing cell cycle progression or to modulate gene transcription. The
products can be used for diagnosis and therapy. They can be used, for
example, to treat tumours or proliferative disorders or spermatogenesis,
osteogenesis, chondrogenesis or the differentiation of progenitor cells.
They can also be used to treat psoriasis, bone diseases,
fibroproliferative disorders, degenerative disorders, or for repair of
cartilage, increasing bone density, liver repair subsequent to a partial
hepatectomy, to promote regeneration of lung tissue in the treatment of
emphysema, or for inducing tolerance in autoimmune diseases, and
transplant recipients. HDx inhibitors can be used as anti-fungal agents,
preservatives in foodstuff, feed supplements for promoting weight gain in
livestock, disinfectants, insecticides or defoliants. The products can
also be used in cell cultures.

Sequence 1449 BP; 406 A; 329 C; 396 G; 318 T; 0 other;

Query Match 89.1%; Score 1434.6; DB 18; Length 1449;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 111 atggcgcaagcgcagggcaccgcggagaaagtcgttactactacgacgggagtgatgga 170
|||||
Db 1 atggcgcaagcgcagggcaccgcggagaaagtcgttactactacgacgggagtgatgga 60
QY 171 aattactattatggacaagggccccaatgaagcctcaccgaatccgcatactcataat 230
|||||
Db 61 aattactattatggacaagggccccaatgaagcctcaccgaatccgcatactcataat 120
QY 231 ttgctgtcaactatgtgtctctaccgaaaaatgaaatctatcgccctcacaaagccaat 290
|||||
Db 121 ttgctgtcaactatgtgtctctaccgaaaaatgaaatctatcgccctcacaaagccaat 180
QY 291 gctgagagatgaccaaggtaccagcgatgactacataaattcttgcgtccatcggt 350
|||||
Db 181 gctgagagatgaccaaggtaccagcgatgactacataaattcttgcgtccatcggt 240
QY 351 ccagataacatgtcggagtagacagcaagcagatgcagagatccaacgttgtagagatgt 410
|||||
Db 241 ccagataacatgtcggagtagacagcaagcagatgcagagatccaacgttgtagagatgt 300
QY 411 ccagtatctgattgctgttctgacgttctgacgttctgacgttctgacgttctgacgtt 470
|||||
Db 301 ccagtatctgattgctgttctgacgttctgacgttctgacgttctgacgttctgacgtt 360
QY 471 gctgtgaaacttaataagcagcagcagacatcgccgtgaattggctggggcgtgcac 530
|||||
Db 361 gctgtgaaacttaataagcagcagcagacatcgctgtgaattggctggggcgtgcac 420
QY 531 catgcaagaagtcgagcgatctgcttctgttaacgtcaatgatatcgtcttgccatc 590
|||||
Db 421 catgcaagaagtcgagcgatctgcttctgttaacgtcaatgatatcgtcttgccatc 480
QY 591 ctggaactgttaagatataccagaggggtgtgtacattgacattgatatccacatggt 650
|||||
Db 481 ctggaactgttaagatataccagaggggtgtgtacattgacattgatatccacatggt 540
QY 651 gacgctgtgaaagagccttctacacacacgacggcggtgacgtgtgtcctttcataag 710
|||||
Db 541 gacgctgtgaaagagccttctacacacacgacggcggtgacgtgtgtcctttcataag 600
QY 711 tatgagaagtagtccaggaactggggacctaagagataccggggtgccaagaagaag 770
|||||
Db 601 tatgagaagtagtccaggaactggggacctaagagataccggggtgccaagaagaag 660
QY 771 tatatgctgttaactaccgctccgagacgggagtgatgacgagtagtctatgagggcat 830
|||||
Db 661 tatatgctgttaactaccgctccgagacgggagtgatgacgagtagtctatgagggcat 720
QY 831 ttcaagccggtcatgttccaaagtaatgagatgttccagcctagtcggtgtcttaacag 890

Db 721 ttcaagccggtcatgtcccaagaagtaatgagatgtccagcctagtgcgtgtcttacag 780
QY 891 tgtgtcagactccctatcttgggatcggttagtltgtctcaatctatctatcaaga 950
Db 781 tgtgtcagactccctatcttgggatcggttagtltgtctcaatctatcaaga 840
QY 951 cagccaagtgtgtggaattgtcaagagctttaacctgtctatgtctgtggagc 1010
Db 841 cagccaagtgtgtggaattgtcaagagctttaacctgtctatgtctgtggagc 900
QY 1011 ggtgttacaccatctglaacgttgcgccgtgtgtgacatatgagacagctgtgccctg 1070
Db 901 ggtgttacaccatctglaacgttgcgccgtgtgtgacatatgagacagctgtgccctg 960
QY 1071 gatacggagatccctaatagtagcttccatacaatgactacttgaatacttggaccagat 1130
Db 961 gatacggagatccctaatagtagcttccatacaatgactacttgaatacttggaccagat 1020
QY 1131 ttcaagctccacatcagtccttccaatatgacttaaccagaacacagaaatgagcttgag 1190
Db 1021 ttcaagctccacatcagtccttccaatatgacttaaccagaacacagaaatgagcttgag 1080
QY 1191 aagatcaaacagcagactgtttgagaaaccttgaatgtctgcgcagcagccctgtggtccaa 1250
Db 1081 aagatcaaacagcagactgtttgagaaaccttgaatgtctgcgcagcagccctgtggtccaa 1140
QY 1251 acgcagcagcttctgtgagagcagccatccctgaaggagagtgcgatgagacgaagacgac 1310
Db 1141 atgcagcagcttctgtgagagcagccatccctgaaggagagtgcgatgagacgaagacgac 1200
QY 1311 cctgacaagcgcactctgactgtctcctctgaacaaacgaattgctgtgaggaagagttc 1370
Db 1201 cctgacaagcgcactctgactgtctcctctgaacaaacgaattgctgtgaggaagagttc 1260
QY 1371 tccgattctgaaggagggagggagggcgcaagaactctccaactcaaaaaaagcc 1430
Db 1261 tccgattctgaaggagggagggagggcgcaagaactctccaactcaaaaaaagcc 1320
QY 1431 aagagaagcaaaaacagagagtgaaaaaagagaagaccagagagagaagaagaatcaccc 1490
Db 1321 aagagaagcaaaaacagagagtgaaaaaagagaagaccagagagagaagaagaatcaccc 1380
QY 1491 gaagaggaagaacccaaggagagagaagccaagccaaggggtcaaggagagggccaag 1550
Db 1381 gaagaggaagaacccaaggagagagaagccaagccaaggggtcaaggagaggggtcaag 1440
QY 1551 ttggcctga 1559
Db 1441 ttggcctga 1449

RESULT 5
AAS81302
ID AAS81302 standard; cDNA; 1582 BP.
XX AAS81302;
AC AAS81302;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #17106.
XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX

PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX
XX (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
PT P-PSDB; ABG17115.
XX

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity

Claim 1; SEQ ID NO 17106; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome CC and gene mapping, and in recombinant production of (II). The CC polynucleotides are also used in diagnostics as expressed sequence tags CC for identifying expressed genes. (I) is useful in gene therapy techniques CC to restore normal activity of (II) or to treat disease states involving CC (II). (II) is useful for generating antibodies against it, detecting or CC quantitating a polypeptide in tissue, as molecular weight markers and as CC a food supplement. (II) and its binding partners are useful in medical CC imaging of sites expressing (II). (I) and (II) are useful for treating CC disorders involving aberrant protein expression or biological activity. CC The polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AAS64197-AAS94564 represent novel human CC diagnostic coding sequences of the invention. CC Note: The sequence data for this patent did not appear in the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1582 BP; 422 A; 376 C; 441 G; 343 T; 0 other;

Query Match 85.5%; Score 1378; DB 23; Length 1582;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1475; Conservative 0; Mismatches 40; Indels 13; Gaps 6;

QY 97 gggagcgagcaagatgycgcagacgagcgaccccggaaggaaagtgttactactacg 156
Db 55 gggagcgagcaagatgycgcagacgagcgaccccggaaggaaagtgttactactacg 114
QY 157 acgggatgttgaaattactattatgtgac-aagccaccaatgaagcctc-accgaa 214
Db 115 acgggatgttgaaattactattatgtgagcgagggcgacccaatgaagcctcgaccga 174
QY 215 ccgatgactcaaatgttctgtcctaactatgtctctaccgaaatggaatctatcg 274
Db 175 ccgatgactcaaatgttctgtcctaactatgtctctaccgaaatggaatctatcg 234
QY 275 ccctacaagaaccaatgtctgagagatgaccaagtaaccacagcgatgactacattaat 334
Db 235 ccctacaagaaccaatgtctgagagatgaccaagtaaccacagcgatgactacattaat 294
QY 335 ctgcgctccatccgctccagataacatgtcgag----tacagcaagcagatgcagaga 389
Db 295 ctgcgctccatccgctccagataacatgtcgaggagtaaccagcaagcagatgcagaga 354
QY 390 ttcaacgttgtagagactgtcc-agtattcgatggcctg-tttagtctgtcagtgtg 446
Db 355 ttcaacgttgtagagacttcccgatattccgatgtgacctgtttagtctgtcagtgtg 414
QY 447 tctactgtgttctgtgcaagtgtctgtg--aaacttaataagcagcagcgacatc 503
Db 415 tctactgtgttctgtgcaagtgtctgtgaaactttaataagcagcagcgacatc 474

QY 504 gccgtgaattgggctggggcctgtcaccatgcaaaagaatccgaggcatctgtctctgt 563
Db 475 gctgtgaattgggtggggcctgtcaccatgtcaaaagaatccgaggcatctgtctctgt 534
QY 564 tacgtcaatgatatactgtcttggccaatccctggaactgtctaagtatacaccagaggtgtctg 623
Db 535 tacgtcaatgatatactgtcttggccaatccctggaactgtctaagtatacaccagaggtgtctg 594
QY 624 tacattgacattgatataatcaccaatgtgtgacggcgtgtgaaagagccctctacaccagcgac 683
Db 595 tacattgacattgatataatcaccaatgtgtgacggcgtgtgaaagagccctctacaccagcgac 654
QY 684 cgggtcatgactgtgtctcttccataaagtatgtgagagtlacttcccaagaaactggggactta 743
Db 655 cgggtcatgactgtgtctcttccataaagtatgtgagagtlacttcccaagaaactggggactta 714
QY 744 cgggataccggggtgtgccaagaacaaagtattatgtctgttaactaccgcgtcccgagacggg 803
Db 715 cgggataatcggggtgtgccaagaagcaagtlattatgtctgttaactaccgcgtcccgagacggg 774
QY 804 attgatgacgagtcctatcaaggccatttccaagccggtcatgtccaagaatgagatg 863
Db 775 attgatgacgagtcctatcaaggccatttccaagccggtcatgtccaagaatgagatg 834
QY 864 ttccagcctagtgcgtgtcttcaagtgtgtgctcagactccctatctgggagatcggtta 923
Db 835 ttccagcctagtgcgtgtcttcaagtgtgtgctcagactccctatctgggagatcggtta 894
QY 924 ggttgcctcaatctatctatcaaaaggacacgcgccaagtgtgtgaaattgtcaagagcttt 983
Db 895 ggttgcctcaatctatctatcaaaaggacacgcgccaagtgtgtgaaattgtcaagagcttt 954
QY 984 aacctgcctatgtctgtatgtctgggagggcggtgtgttacaccattcgttaacgtttgccggtgc 1043
Db 955 aacctgcctatgtctgtatgtctgggagggcggtgtgttacaccattcgttaacgtttgccggtgc 1014
QY 1044 tggacatatgagacagctgtgtgcccctgtgatacggagagatccctaatagtagcttccataaat 1103
Db 1015 tggacatatgagacagctgtgtgcccctgtgatacggagagatccctaatagtagcttccataaat 1074
QY 1104 gactacttgaataacttggaccagaatttcaagctccacatcagtccttccaataatgact 1163
Db 1075 gactacttgaataacttggaccagaatttcaagctccacatcagtccttccaataatgact 1134
QY 1164 aaccagaacacgaatgagtaactgtgagaagaatcaaacacgagactgtttgagaacctaga 1223
Db 1135 aaccagaacacgaatgagtaactgtgagaagaatcaaacacgagactgtttgagaacctaga 1194
QY 1224 atgtgtccgacagcacctgtgggttccaacgacgagcgatctcctgagagcgccatccctgag 1283
Db 1195 atgtgtccgacagcacctgtgggttccaacgagcgatctcctgagagcgccatccctgag 1254
QY 1284 gagagtggcgatgagagcagaagacgcacctgacaagcgcatctcgatctgcctctgac 1343
Db 1255 gagagtggcgatgagagcagaagacgcacctgacaagcgcatctcgatctgcctctgac 1314
QY 1344 aaacgaattgtcgtgtgagaagaagtlctccgatlcttaagaaggagaggagggcgccgc 1403
Db 1315 aaacgaattgtcgtgtgagaagaagtlctccgatlcttaagaaggagaggagggcgccgc 1374
QY 1404 aagaactcttccaacttcaaaaaagccaagagagtcataaacagagatgataaaagagaaa 1463
Db 1375 aagaactcttccaacttcaaaaaagccaagagagtcataaacagagatgataaaagagaaa 1434
QY 1464 gaccacagagagaagaagaatcacccaagagaagaaaccaaggaagagaagccaagaa 1523
Db 1435 gaccacagagagaagaagaagtcacccaagagaagaaaccaaggaagagaagccaagaa 1494
QY 1524 gccaaaagggttcaaggaaggccaagtgtgcctgtaatgacctctccagctctgtcttc 1583
Db 1495 gccaaaagggttcaaggaaggccaagtgtgcctgtaatgacctctccagctctgtcttc 1554

QY 1584 ctgtctagtcctcacggttcttcccc 1611
Db 1555 ctgtctagtcctcacggttcttcccc 1582

RESULT 6
ABI99512
ID ABI99512 standard; cDNA; 1997 BP.

AC ABI99512;
XX
DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:493.

KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia; vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

OS Mus musculus.

PN WO200188188-A2.

PD 22-NOV-2001.

PF 18-MAY-2001; 2001WO-JP04192.

PR 18-MAY-2000; 2000JP-0145977.

PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

DR WPI; 2002-034733/04.

DR P-PSDB; ABB57202.

PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring expression levels of particular genes defined in the specification or by determining the expression profile of a gene group comprising these genes -

PS Claim 2; Page 1369-1373; 2690pp; English.

XX The present invention describes a method for examining ischaemic conditions, comprising measuring the expression levels of particular genes (I) in a test sample or determining the expression profile of a gene group in the sample comprising genes selected from (I). The method is useful for examining the ischaemic condition (e.g. compressive ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring expression levels of particular genes (ABI99202 to ABI99912, encoding the protein sequences in ABB57020 to ABB57374) or by determining the expression profile of a gene group comprising these genes. The expression levels or expression profiles produced by these genes are used as an indicator when screening for ischaemic condition-improving drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR primers for a mouse ischaemic condition related sequence, which are used in the exemplification of the present invention.

SQ Sequence 1997 BP; 595 A; 376 C; 486 G; 540 T; 0 other;

Query Match 50.8%; Score 818; DB 24; Length 1997;
Best Local Similarity 73.4%; Pred. No. 3.8e-238;
Matches 1046; Conservative 0; Mismatches 380; Indels 0; Gaps 0;

QY 126 ggcacccggaggaagtgtgtactactacgacggygagttggaattactattatga 185

Db 187 ggcggcaagaagaagtgtgtactactatgatgtgtgatatgtgcaattattatgyc 246

QY 186 caagccaccaatgaagctcacccgaatccgcatgataataattgtgtctcaactat 245

Db 247 cagggtcatcccatgaagcctcatagaatccgcatgatacataactgtgtctaaattat 306

QY 246 ggtctctacccaanaaatgtgaatatctatcgccctcacaaagccaatgtgtgagatgacc 305

Db 307 ggttataccgaaaaatggaatatatatagcgctcataaagccactgctgaaagaatgact 366
QY 306 aagtaaccacagcgatgactatcatatattcttgcgctccatccgctccagataacatgctg 365
Db 367 aaataccacagcgatgataatcataagttcttaacgatcaataagaccagataatagtct 426
QY 366 gagtaacagcaagcagatcagagagattcaacggttgtaggaagactgtccagtatctgatg 425
Db 427 gagtaacagtaagcagatgcaagagatttaacgctcgagaagattgtccggtgttgatgga 486
QY 426 ctgtttgagttctgtcagttgtctactggtgtgtctgtggaagtgctgtgaaacttaac 485
Db 487 ctcttgagtttgcagctctccacggtgtgtcagttgtggtgctggtggaattaaac 546
QY 486 aagcagcagacgacatccgctgtaattgggctggggcctgcaccatgcaagaagtcc 545
Db 547 cggcaacaactgataatgctgtcaatttggtgaggaactacatcatgccaagaagtca 606
QY 546 gaagcatctggtctgttaacgtcaatgatatcgcttgccatccctggaactgctaag 605
Db 607 gaagcatcagggttctgtcattgttaatatgtgtctgccaatccctgaattacttaag 666
QY 606 tatcaaccagaggtgtgtacattgacattgatatccaccatggtgacggcgctggaagag 665
Db 667 tatcatcagagagctctatatattgacataagacatccaccatggtgattgttgaggaa 726
QY 666 gcctctacacacacgacggcggtcatgactgtgtcctttcataagtatgagagtaacttc 725
Db 727 gcttttatacaacacagatcggtgatgacggctcattccataataatgaggaaacttt 786
QY 726 ccaggaactggggacctaaccgggagataccgggctgcaaaagacaagatlatgctgttaac 785
Db 787 cctggaacaggaacttgaggagataattggtgtgtaaaaggaaataactatgtgtcaat 846
QY 786 taaccgctccgagacggagattgtagcagagtcctatgaggccatttcaagccggtcatg 845
Db 847 ttcccatgagagatgtagatgtagatgataatgacataatgacaattttaaagcctatcatc 906
QY 846 tccaaagtaatgagatgttccagcctagtcggtgtgtcttacagtgtgctcagactcc 905
Db 907 tcaaaagtatgtagatgtagcagcctagcgcgtggtgtgctgagtggtgcgagactcc 966
QY 906 ctatctgggagtcggttaagttgtgttcaatctatcatcaaaagacacgccaagtgtgtg 965
Db 967 ctgtctgggacagcgctggtgtgttcaatctaactgaactgcaaaagtcagtctaattgtgta 1026
QY 966 gaattgtcaagagctttaacctgacctatgtgtgagtgaggcggtgtgttaaccatt 1025
Db 1027 gaagtagcgaaaactttaacttgccattgtgctgagtcgctgtaggagggcttacaacatc 1086
QY 1026 cgtaacgttgcccggtgtgtgacataatgagacagctgtgcccctgatacggagatccct 1085
Db 1087 cggaatgttgcccgatgttgacataatgagactgcagttgccccttgattgtgaattccc 1146
QY 1086 aatgagcttccatacaatgacttcttgaaactttgaccagatttcaagctccacatc 1145
Db 1147 aatgagttgccataataatgattactttgagattttgaccagacttcaactgcatalt 1206
QY 1146 agtccttccaatatgactaacacagacaagaaatgtagtacctgtagaagatcaaacagcga 1205
Db 1207 agtccttccaacatgacaacacagacactccagaatatatgaaagataaacaagcgt 1266
QY 1206 ctgttgagaaccttagaatgtgtccgcacgcacgtgggtccaacgcagcgattcct 1265
Db 1267 ttattgaaaatctacgtatgttaccacatgcacctgtgttcaaatgtcaagctattcca 1326
QY 1266 gaggaagccatccctgtaggagagtggtgagatgagacgaagacccctgacaagcgcatc 1325
Db 1327 gaggaatgctgttcatgaaagacagtgtagatgagagatggaagaagacccggacaagaat 1386
QY 1326 tcgatctgtcctctgacaacaagaaatgtcgctgtgaggaagagttctccgatctgaaagag 1385

Db 1387 tccattcgagcatcagacaacaacgtagatgttgcgatgaaagagtttccagattctgagat 1446
QY 1386 gagggagagggggccgcaagaactcttccaacttcaaaaaagccaagagagtcanaaca 1445
Db 1447 gaagtgaaaggaggtcgttagaatgttgcgtgatacagaagagagcaagaagagctag 1506
QY 1446 gagatgaaaaagagaaagaccacagagagagaagaaagaaatcacccgaagaggagaaacc 1505
Db 1507 attgaagaagacaagaagagagacagagagcacaagaagacagatgtttaagaaagacaaa 1566
QY 1506 aaggagagaagccgaagccaaggggtcaaggagagagccaagt 1551
Db 1567 tccaaggaacaatagtgtgtagaanaaacagacaccccaaggaagccaagt 1612
RESULT 7
AAZ32066
ID AAZ32066 standard; DNA; 1985 BP.
XX
AC AAZ32066;
XX
DT 10-JAN-2000 (first entry)
XX
DE Human METH2 related EST U31814.
XX
KW Human; METH1; METH2; anti-angiogenic; metalloproteinase thrombospondin;
cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
angiogenesis inhibitor; abnormal wound healing; inflammation;
rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
diabetic retinopathy; macula degeneration; haemangioma; detection;
arterial-venous malformation; immune deficiency; ss.
XX
OS Homo sapiens.
XX
PN WO9937660-A1.
XX
PD 29-JUL-1999.
XX
PF 22-JAN-1999; 99WO-US01313.
XX
PR 23-JAN-1998; 98US-0072298.
PR 28-AUG-1998; 98US-0098539.
XX
PA (IRUE/) IRUELA-ARISPE L.
PA (HAST/) HASTINGS G A.
PA (RUBE/) RUBEN S M.
PI Irue-la-Arispe L, Hastings GA, Ruben SM;
XX
DR WPI; 1999-590684/50.
XX
PT New isolated metalloproteinase thrombospondin polypeptides, useful for
PT treating hyperproliferative disorders, cancers or autoimmune disorders
PT
XX
PS Disclosure; Page 444-446; 457pp; English.
XX
CC AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC metalloproteinase thrombospondin (METH) proteins METH1 and METH2
CC respectively. METH1 and METH2 have been found to be potent inhibitors of
CC angiogenesis both in vitro and in vivo. They can be used for treating
CC cancer and other disorders related to angiogenesis including abnormal
CC wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC endometrial bleeding disorders, diabetic retinopathy, some forms of
CC macula degeneration, haemangiomas, and arterial-venous malformations.
CC They may be useful in treating deficiencies or disorders of the immune
CC system, by activating or inhibiting the proliferation, differentiation,
CC or mobilisation (chemotaxis) of immune cells. The etiology of these
CC immune deficiencies or disorders may be genetic, somatic, such as
CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC toxins), or infectious. They can also be used to treat inflammatory
CC conditions, both chronic and acute conditions. The products can also be
CC used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to

CC AAY49511 represent sequences given in the exemplification of the present
invention.
XX
SQ Sequence 1985 BP; 626 A; 360 C; 454 G; 545 T; 0 other;

Query Match 50.0%; Score 805.2; DB 20; Length 1985;
Best Local Similarity 72.8%; Pred. No. 3e-234;
Matches 1038; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

QY 126 ggcacccggaggaaagtcgttactactacgacgggagatgttggaaattactattatgga 185
Db 223 ggcgcgcaaaaaaagctgtactactacgacggtgatatttggaaattattattatgga 282
QY 186 caagggcaccacaatgaagcctcacccgcatgcatgactcataatttgcgtcaactat 245
Db 283 cagggtcatcccatgaagcctcatagaaatccgcatgaccataacttgcgttaaatat 342
QY 246 ggtcctacgcgaaaaatggaatctatgcctcacaaagccaatgctgagagatgacc 305
Db 343 ggccttataccgaaaaaatggaatatatagcccataaagccaactgcggaagaatgaca 402
QY 306 aagttaccacagcgatgactacataatcttctgcgtccatccgtccagataacatgtcg 365
Db 403 aaatatcacagtgatgagtataatcaaatcttctacggtlcaataaagaccagataacatgtct 462
QY 366 gagtacacgacgacagatgacagagattcaacglttgtagagactgtccagtattcgatgac 425
Db 463 gagtatagtgaagcagatgacataataatttggagaagattgtccagcglttgatgga 522
QY 426 ctgtttgagtcctgtcagttgttactactggtgttctgtggaagtgctgtgaaactaat 485
Db 523 ctcttgagttgttcagctctcaactgycggttcagttgtgagcgtgtggaaglttaaac 582
QY 486 aagcagcagacgacacatgcgcgtgaattggcgctgggacctgcaccatgcaaaagaagtc 545
Db 583 cgacaacagacgatgatalgctgttaattggcgctggagattacatcatgctaaagaatac 642
QY 546 gaggcacatggtcttctgttacgtcaatgatatcgttctggccatccctggaactgctaaag 605
Db 643 gaagcatcagagattctgttacgttaattggtgagattacatcatgctgaattactaaag 702
QY 606 tatcacccagaggggtgtgtacattgacattgatatccaccatgtgtgacggcggtggaagag 665
Db 703 tatcatcagagagctcttatatatattgatatagatattcatcatggtgagtgttgaaaga 762
QY 666 gccttctacaacacggaacggggtcatgactgttccttccataagtatggagagtacttc 725
Db 763 gctttttatacaacagatcgtgttaactgacggtatcatccataaataatggggaatacttc 822
QY 726 ccagggaacttggggaacctacgagataccggggtcggcaaaagacagatatatgctgttaac 785
Db 823 cctggcacagagagacttggaggatattggtgtcggaaaaaggaataatactatgctgaat 882
QY 786 taaccgctccgagacgagatgtatgacaggtccctatgagccatttcaagccggtcatg 845
Db 883 ttccaatgtgtgatatgatatgtagtgcataatggtgcagatatattaagcctattatc 942
QY 846 tccaagtgtaatggagatgttccagcctagtcggtgtcttaccagtgtgctcagactcc 905
Db 943 tcaaaagtgatggagatgtatcaaccctagtgtgtgtattaccagtggtgtgcagactca 1002
QY 906 ctatctgggagatcgtttagttgttcaatctatctatcaaaaggacacgccaagtgtgtg 965
Db 1003 ttatctgtgtagtagactggtgttcttcaatctacaacagtcacaaggtcactaatgtgta 1062
QY 966 gaatttgcacaagactttaacctgcctatgctgtagtggagggcggtgttaccacatt 1025
Db 1063 gaagttgttaaaaaactttaacttaccattactgtagtcttggaaggaggtgtctacacaatc 1122
QY 1026 cgtaacgttcccggtgtgtgacatatgagacagctgtggccctgatacggagatccct 1085
Db 1123 cgtaatgtgtcgtcgtatgttgacatatgagactgcagltgccccttgatgtgagattccc 1182

QY 1086 aatgagcttcacatacaatgactactttgaatactttggaccagatttcaagctccacatc 1145
Db 1183 aatgagttgcacataaatgattacttlttgatattttggaccagacttcaactgcatatt 1242
QY 1146 agtccctccaatatgactaaccagaacagaaatgagtlaccttggagaagatcaaacagcga 1205
Db 1243 agtccctcaaaatgacaaacaccagacacccccaagatatatggaagaagataaaacagcgt 1302
QY 1206 ctgtttgagaaccttagaatgtgtccgcgacacacctggggtcccaacgcaaggcatttcc 1265
Db 1303 ttgtttgaaaatttgcgcattgttaccttcacacctgtgtgtccagatgcaagctattcca 1362
QY 1266 gaggaacgcatccctgaggaagatggcgatgagacggaagacgaccctgacaagcgcatc 1325
Db 1363 gaagatgtgttcatgtaagacagltggagatgaagatggagaagatccagacaagaagatt 1422
QY 1326 tcgattgtctcctctgacaaacgaattgctgttgaggaagagttctcgattctgaagag 1385
Db 1423 tctattcgagatcagacaacgagatagctgtgtgatgaagaatttccagattctgagat 1482
QY 1386 gagggagagggggcccgaagaactcttccaacttcaaaaaagccaagagagtlcaaaaca 1445
Db 1483 gaaggagaaggaggtcgaagaatgtgtgcgtatcataagaagaaggcaagaagctaga 1542
QY 1446 gagatgaaaaaagaaagacccagagggagaagaagaatccaccgaaagaggaanaacc 1505
Db 1543 attgaagaagaataagaagaacacagagagcaaaaaaacagacglttaaggaaagataaa 1602
QY 1506 aaggagagaagccagaaagccaaagggtccaaggagagggccaagt 1551
Db 1603 tccaaggacacagtggtgtaaaaaacagataccaagaagcaacaat 1648

RESULT 8
AAC90323
ID AAC90323 standard; DNA; 1985 BP.
XX
AC AAC90323;
XX
DT 19-MAR-2001 (first entry)
XX
DE U31814 cDNA clone.
XX
KW METH; metalloprotease; thrombospondin; angiogenesis inhibition;
cancer therapy; benign tumour; ocular angiogenic disease;
rheumatoid arthritis; psoriasis; wound healing; endometriosis;
vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
scleroderma; trachoma; vascular adhesion; myocardial angiogenesis;
coronary collateral; cerebral collateral; arteriovenous malformation;
ischaemic limb angiogenesis; Osler-Webber syndrome; wound granulation;
plaque neovascularisation; telangiectasia; haemophilic joint; EST;
angiofibroma; fibromuscular dysplasia; expressed sequence tag;
Crohn's disease; atherosclerosis; birth control; ss.
XX
OS Unidentified.
XX
PN WO200071577-A1.
XX
PD 30-NOV-2000.
XX
PF 25-MAY-2000; 2000WO-US14462.
XX
PR 25-MAY-1999; 99US-0318208.
PR 20-JUL-1999; 99US-0144882.
PR 10-AUG-1999; 99US-0147823.
PR 13-AUG-1999; 99US-0373658.
PR 22-DEC-1999; 99US-0171503.
PR 22-FEB-2000; 2000US-0183792.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

PA (IRUE/) IRUELA-ARISPE L.
PA (HAST/) HASTINGS G A.
PA (RUBE/) RUBEN S M.
PA (JONA/) JONAK Z L.
PA (TRUL/) TRULLI S H.
PA (FORN/) FORNWALD J A.
PA (TERR/) TERRETT J A.
XX
PI Irueula-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;
PI Fornwald JA, Terrett JA;
XX
XX WPI: 2001-025136/03.

PT METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT inhibit angiogenesis in the treatment of disorders such as cancer,
PT rheumatoid arthritis and psoriasis -

XX Claim 14, Pages 745-746; 768pp; English.

XX The present invention relates to human METH1 and METH2, (ME for
CC metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
CC The present sequence and TH is an expressed sequence tag (EST) for METH. METH
CC can be used for inhibiting angiogenesis in an individual, and for
CC treating cancer, benign tumours, an ocular angiogenic disease,
CC rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
CC scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
CC coronary collaterals, cerebral collaterals, arteriovenous malformations,
CC ischaemic limb angiogenesis, Osler-Weber syndrome, plaque
CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
CC fibromuscular dysplasia, wound granulation, Crohn's disease or
CC atherosclerosis. METH can also be used in birth control. METH can also
CC be used in diagnostic methods for the prognosis of cancer.

SQ Sequence 1985 BP; 626 A; 360 C; 454 G; 545 T; 0 other;

Query Match 50.0%; Score 805.2; DB 22; Length 1985;
Best Local Similarity 72.8%; Pred. No. 3e-234;
Matches 1038; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

QY 126 ggcacccgaggaagtctgttactactacgacgggagtgtggaattactattatgga 185
DB 223 ggcggcaaaaaaagtctgttactactacgacgggtgatattggaattattatgga 282
QY 186 caagggcaccatgaagccctcaccgaatccgcatgactataattgtgtccaactat 245
DB 283 cagggtcatcccatgaagccctcaccgaatccgcatgaccataactgtgtlaaattat 342
QY 246 ggtctctaccgaaaaatggaatactatcgccctcacaagaacaaatgtgagagatgacc 305
DB 343 ggcctatacagaanaatggaataatagagcccaataagccaactgccaagaanaatgaca 402
QY 306 aagtaaccacagcagatgactaataattctgtgcgtccatccgtccagataacatgctcg 365
DB 403 aaatacacagtgatgagtaataaatttctacggtcaataaagaccagataacatgctct 462
QY 366 gagtacagcaagcagatgcagagatcaacggtgtgagagactgtccagtatccgagtcgc 425
DB 463 gagtatagtaagcagatgcagatataatttctacggtcaataaagaccagataacatgct 522
QY 426 ctgtttgagttctgtcagttgttactgtgtgttctgtggaagtgtgtgaaactaat 485
DB 523 ctctttgagttttgtcagctctcgaactgcggttcagttcgtgtgagctgtgaaagttaac 582
QY 486 aagcagcagacgagacatgcgcgtgaattggcgtggggcctgcaccatgcaagaagtcc 545
DB 583 cgacaacagactgatatgctgttaattggtggtgaggaattacatcatgtctaagaatac 642
QY 546 gaggcactctgctctgttactgtaaatgatatcgctcttgccatccctggaactgctaag 605
DB 643 gaagcatcagatctgttactgtaaatgatatgtgtgtgccaatccctgaattactaaag 702

QY 606 tatcaccagagggtgctgtacattgacatgtatatccaccatggtgacgsgtggaaag 665
DB 703 tatcatcagagagctcttatatatgtatatagatatcatcatggtgatgtgtgaaaga 762
QY 666 ggcctctacaccacgagccgggtcatgactgtgtcccttcataagtatggaagtacttc 725
DB 763 gcttttatcacacagatcgtgtaatgacggtatcatcattccataaatatgsggaactt 822
QY 726 ccaaggaactgggacctaagcgataccgggctggcaagaacaagtattatctgttaac 785
DB 823 cctggcacagagagacttgagggatattgtgtctggaaaaagcaataactatgctgcaat 882
QY 786 taccgctccgagacgggattgtacgagtcctatgagccatttccaagccggtcatg 845
DB 883 ttccaatgtgtgatgtatagatgatgagtcataatgsgcagatatttaagcctattatc 942
QY 846 tccaaagttaatggagatgttccagcctagtcggtgtgtcttacagtggtgctcagaactc 905
DB 943 tcaaaagtgatggagatgtatcaaccctagtcgtgtgtattacagtggtgtgagactca 1002
QY 906 ctatctggygacgtgttaagttgtcttcaatctatctatacaaaaggacacgccaagtgtg 965
DB 1003 ttatctgtgatagactggtgtgttccaatctaacagtcaaaaggtcatgctaagtgtga 1062
QY 966 gaattgtcaagagcttcaacctgctatgctgacgtgtggaagcggtgtgtacaccatt 1025
DB 1063 gaagtgtaaaaaacttcaacttaccattactgacgtgtggaaggtgtgccaacaatc 1122
QY 1026 cgtaacgttgcccggtgtgcagacatatgagacagcgtgtggccctgatacagagatccct 1085
DB 1123 cgtaatgtgtcgtcgtatgttgacatatgagacactgacgttgccctgattgtgagattccc 1182
QY 1086 aatgagcttcacatacagactacttctgaatactttggaccagatttcaagctccacatc 1145
DB 1183 aatgagttgcacataataatgattacttcttgagttatttggaccagacttcaactgtcatatt 1242
QY 1146 agtccttccaatatgacttaaccagaacaacagaaatgagtaaccttggaagaatcaaacagcga 1205
DB 1243 agtccttccaatatgacttaaccagaacaacagaaatgagtaaccttggaagaatcaaacagcgt 1302
QY 1206 ctgtttggaacaccttagaatgtgcgcgacgacactggygtccaacgcaagcgatctct 1265
DB 1303 ttgtttgaaatttgcgatgttaccctcatgacactggtgtccagatgcaagctattcca 1362
QY 1266 gaggaagccatccctgagagagtggtgagatgagagcaagacccctgacaagcgcatc 1325
DB 1363 gaagatgctgttcatgtaagacagtgtagatgagaatggaagatccagacaagaagaatt 1422
QY 1326 tcgatctgtctcctctgacaacgaattgtccctgtgtggaagaagttctccgattctgaagag 1385
DB 1423 tctattcgagcatcagacaagcggatagcttgtgtgaagaattctcagattcttgagat 1482
QY 1386 gagggagagggggccgcaagaactcttccaacttcaaaaaaagccaagagagtcacaaca 1445
DB 1483 gaaggaagaagaggtcgaaagaatgtgtgctgatacataaagaagaagcaagaagctaga 1542
QY 1446 gaggaatgaaaaaagaagaaccagaaggaagaagaagaatcaacgaagaaggagaacc 1505
DB 1543 attgaagaagataagaagaagaacagagacaaaaaacaagacgttaaggaagaagataaa 1602
QY 1506 aaggaaggaagccagaagccaaggggtcaaggaaggagccaagt 1551
DB 1603 tccaaggaacaacagtggtgaaaaaacaagatatccaagaagaaaccaaat 1648

RESULT 9
AAC89555
ID AAC89555 standard; DNA; 1985 BP.
XX
AC AAC89555;
XX
DT 08-MAR-2001 (first entry)
XX

DT	26-MAR-2002 (first entry)	
XX		
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 173.	
XX		
KW	Drosophila; developmental biology; cell signalling; insecticide;	
KW	pharmaceutical; gene; ss.	
XX		
OS	Drosophila melanogaster.	
XX		
PN	WO200171042-A2.	
XX		
PD	27-SEP-2001.	
XX		
PF	23-MAR-2001; 2001WO-US09231.	
XX		
PR	23-MAR-2000; 2000US-191637P.	
PR	11-JUL-2000; 2000US-0614150.	
XX		
PA	(PEKE) PE CORP NY.	
XX		
PI	Venter JC, Adams M, Li PWD, Myers EW;	
XX		
DR	WPI; 2001-656860/75.	
DR	P-PSDB; ABB57794.	
XX		
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -	
PT		
XX		
PS	Claim 1; SEQ ID NO 173; 21pp + Sequence Listing; English.	
XX		
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins	
CC	(ABB57737-ABB72072).	
CC	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.	
CC		
XX		
SO	Sequence 2145 BP; 586 A; 550 C; 556 G; 453 T; 0 other;	

Query Match 45.3%; Score 729; DB 23; length 2145;
Best Local Similarity 73.1%; Pred. NO. 5e-211;
Matches 936; Conservative 0; Mismatches 345; Indels 0; Gaps 0;

QY	488	gcagcagacgacatcgccgtgaattgggctggggcctgcaccatgcacaagaagtcgca	547
Db	595	gcaagcctcgagatctgcatacaattggggcggtggtgcatacagcccaaaaatcgca	654
QY	548	ggcatctgctctgttaacgtcaatgatatacgtcttggccatcctggaactgctaagta	607
Db	655	agcctcggtctctgtcaagcaacgacatgttcttgggaattcttgaactgtctaata	714
QY	608	tcaccagagggtgtgtacattgacattgatatccacatggtgaacgctggaagagcg	667
Db	715	ccatcagcgtgtctctacatagatatagacgtccatcaacgycgatggcgtggaagagc	774
QY	668	cttctacaccacgaccgggtcatagactgtgtcctttcataaagtatggaagtaactccc	727
Db	775	gtctataccaccgatcgtgtgatgtactgtcagcttccacaagtaacgagagatattccc	834
QY	728	aggaactggggaccctacgggataccgggctggcaagaagtattatgtgttaacta	787
Db	835	gggcactggcgatctgcgagacattggcgccggcaagggaagtactatgcggtgaatat	894
QY	788	cccgctccgagacgggattgatgacaggtcctatgaaggccatttcaagccggtcatgtc	847
Db	895	accctgcgtgatgcatgatgatgatgacgtcgtacgaagacatatattgtgccattatcag	954
QY	848	caaaatgatgagatgttccagcctagtgcggtgtgtcttacagtggtgtcctaagactccct	907
Db	955	caaggtgatggaacattccagccggcagccggtgtgtgtgcagtggtggcccgattcgct	1014
QY	908	atctgggatacggttaggtgtgtcttcaatctatctatacaaaaggacacgccaagtgttgya	967
Db	1015	gactggcgatcgtgttagtgcgtcttcaatctcaccgctcaagggtcacgcaagtgtcgtga	1074
QY	968	attgtcaagagctttaacctgcctatgctgatgctggagcggtgtgttacacaacttcg	1027
Db	1075	gttcgtgaagaataataacctgcattcctctgatgtgcgcggtgtgtgtataccaactcg	1134
QY	1028	taacgttgcgcggtgtgcacatatgagacagcgtgtggccctgagatacagagatccctaa	1087
Db	1135	taatgtatcccgctgcggacattatgagacctccgttgcaactggccgtgagatagccaa	1194
QY	1088	tgagcttcatacaatgactacttltgaatactttggaccagatttcaagctccacatcag	1147
Db	1195	cgaactgcctacaacgacttacttgcagtaactttggcccgattttaaagctgcacattag	1254
QY	1148	tccttccaatatgacttaaccagaacacgaatgatactgtgagaagatcaaacagcgact	1207
Db	1255	tcccagcaatatgacgaatcagaataacatccgagtaacctggagaagatcaagaaccgltc	1314
QY	1208	gtttggaaccttagaatgtcgcgcgaacgcaacttgggtgtccaaacgcagcgatctctga	1267
Db	1315	gttcgagaacctgacgatgtcgtcctaagctccgggcttcaaatccaagcgattccgga	1374
QY	1268	ggaagccatccctgagagagagtgatgagagcaagaacacccctgacaagcgcatctc	1327
Db	1375	ggaatgcatacaacgatacgttcgcgacgagagcaaggttcgacaagagatgcgcctgc	1434
QY	1328	gactgtctcctctgacaacgaatgtgcctgttgaggaagagttctccgattctgaagaga	1387
Db	1435	gcagagcgacaaggaacgcatgtgtgccggaacgagtaactccgattcgagagatga	1494
QY	1388	gggagaggggggccgcaagaa 1408	
Db	1495	gggcgaagcggtcgcaaggga 1515	

RESULT 11
ABL01896/C
ID ABL01896 standard; cDNA; 4755 BP.
XX ABL01896;
AC
XX
DT 26-MAR-2002 (first entry)

[illegible]

OY	626	catltagacattgatatattcaaccatgtgtacagcgctlygaagaaggcccttcacaccacgygaccg	685
Db	2508	CATGATATATAGACGGTCCATCATCAGCGGCATGGCGTTGAGAGAGGGCGTTCTATTACCACCGATCG	2449
OY	686	ggtcactgactgtgtccttccataagtatgtagagagtlactcccagaactgyggacctagc	745
Db	2448	TGTGATGACTGTCAAGCTTCACACAAGTAGCAGAGATATTTCCCGGGCACTGGCGATCTGCCG	2389
OY	746	ggtataccggggctgycgacaagaagiatatatgtgttaactaccgcgtccgagacygat	805
Db	2388	AGACATTTGGGGCCCCGGAAGAAAGTAATAAGCGGTGATATATACCCCTGGCGTGATGGCAT	2329
OY	806	tgtatgacgagtgctcctatgtaggccattttccaagccggtcatgtgccaaqtaatgagatgtt	865
Db	2328	GGATGATGATGGCGTAGAGAGACATATTGTGCCCATTTATCAGCAAGGTGATGGAACAATT	2269
OY	866	ccagccctaagtcggtgtgtccttacagtgtygtctcagactcccatactcygggatcgtgtlag	925
Db	2268	CCAGCGCGGACGCCGTGTTGTCAGATGTGGCGCCGCAATTCCGCTGACTGGCGATCGGTTAGG	2209
OY	926	ttgcttcaatctatctatcaaagaacacgccaagtgtgtggaattgtccaagagcttta	985
Db	2208	CTGCTTCAATCTCACCGTCAAGGGTCAAGGCCAAGTGCGTGGAGTTCTGAAGAATATATA	2149
OY	986	cctgcctaagtctgatlgtcgtggaagcgytgglttaacaacattcgtlaacgttgcgcgtgtcgt	1045
Db	2148	CCTGCATTCCTGATGTTGCGCGGTGGTTATACCATTTCGTAATGTATCCCGCTGCTG	2089
OY	1046	gacatatgagacagcgtgtgccccgtgatacgcgagatccctaattgagattccataaatga	1105
Db	2088	GACCTATGAGACCTCCGTTGCACCTGGCGGTGAGATAGCCAAGSACTGCCCTACACAAGA	2029
OY	1106	ctactttgataactttgagaccagatttcaagctccacatcagtccttccaatatgactaa	1165
Db	2028	TTACTTTCGACTACTTTTGGGCCGATTTTAAGCTGCACATTAGTCCCAGCAATATGACGAA	1969
OY	1166	ccagaacacgyaatgagtlactctgtagaagatcaaacagcgcactgttttgaagaaccttagaat	1225
Db	1968	TCAGAAATACATCCGAGTACCTGGAGAGATCAAGAACCCTGTCTGTCGAGAACCCTGGCGAT	1909
OY	1226	gctgcgcgacgcacactggyggtccaaaagcagggcgtattcctgagagagccatccctgaaga	1285
Db	1908	GCTGCCCTCACGCTCCGGCGTTCAAAATCCAAAGCGATTCCCGAGGATGCCATCAACGATGA	1849
OY	1286	gagtgcgatgagagacgaagacgaccctgacaagcgcacatctgacgtctcctctgacaa	1345
Db	1848	GTCGACGACGAGGACCAAGGTCGACAAAGATGATCGCTGCGCGCAGACGACAAGACAA	1789
OY	1346	acgaattgcctgtgaggaagatctccgatcttgaagagagagagagggggcgccgcaa	1405
Db	1788	GCGCATTTGTCCCGAGAACGAGTACTCGGATTCGAGAGATGAGGGCGAAGGGCGTGCGCAG	1729
OY	1406	gaa 1408	
Db	1728	GGA 1726	
RESULT 12 AAAX90840 ID AAAX90840 standard; DNA; 1943 BP. XX AC AAAX90840; DT 13-JAN-2000 (first entry) XX DE Maize histone deacetylase-4 DNA. XX KW Maize histone deacetylase; family 1, ZmHDI; promoter regulator; promoter; KW RNA polymerase II; transcription; plant transformation; heterochromatin; KW disease resistance; chromatin assembly; gene activity; toxin screening; KW pathogenicity; disease response promoter; ds. XX OS Zea mays.			

XX	Key	Location/Qualifiers
EH	CDS	51..1610
FT		/*tag= a
FT		/product= "Maize histone deacetylase"
XX		WO9951731-A2.
PN		14-OCT-1999.
XX		02-APR-1999; 99WO-US07370.
XX		03-APR-1998; 98US-0080563.
PR		(PION-) PIONEER HI-BRED INT INC.
XX		Baldwin DA, Briggs SP, Crane VC;
XX		WPI; 1999-611038/52.
DR		P-PSDB; AAY28800.
XX		New deacetylase genes, used for producing transgenic plants which have
PT		increased disease resistance
XX		Claim 1; Page 60-63; 87pp; English.
XX		The present sequence encodes a maize histone deacetylase. This DNA
CC		belongs to family 1, ZmHD1 and appears to be regulator of promoters for
CC		RNA polymerase II, for transcription of genes. The nucleotide sequence
CC		can be used to transform plants and increase disease resistance by
CC		altering the levels of heterochromatin, altering chromatin assembly, and
CC		gene activity of the transformed plants. Additionally, compositions find
CC		use in screening for toxins that affect pathogenicity and in determining
CC		which disease response promoters are regulated by histone deacetylase.
XX		Sequence 1943 BP; 510 A; 498 C; 497 G; 438 T; 0 other;
SO		
QY	Query Match	31.2%; Score 502.6; DB 20; Length 1943;
	Best Local Similarity	65.5%; Pred. No. 4.1e-142;
	Matches 768; Conservative	0; Mismatches 399; Indels 6; Gaps 2;
QY	134	gaggaaagctgttactactacagcgggagtgttgaaattactattatgacaaggcca 193
DB	122	gcggcgctgtgtacttctacgaccggagtgtggcaactactactacgggcaaggcca 181
QY	194	cccaatgaagcctcaccgaatccgcatgactcataattgtctgctcaactatgtctcta 253
DB	182	tccgatgaagcgcgcgcatcccgatgacgactcgtctgtgcgcgtactagggcctcct 241
QY	254	ccgaaaaatggaatctatcgcgcctcacaagaaccaatgctgagagatgaccaagtacca 313
DB	242	caaccagatgcagtgtaacgcccacacccgcgcgcgacgcgacgcactctgcgcttcca 301
QY	314	cagcgatgactacatnaattcttgcgctccatccgctccagataacatgtcggagtacag 373
DB	302	cgccgacgactacatcaacttctgtcgctccgtccgacgcgcgaaacgcagcagaccagat 361
QY	374	caagcagatgcagagattcaacgctgtgtgagagactgtccagtatctcgatgcgctttga 433
DB	362	ccgctgtctcaagcgttcaacgctgcgcgagagactgcccgtcttcgacggcctctacag 421
QY	434	gttctgtcagttgtactgtgtgtctgtgtgcaagtgctgtgaaacttaagaacagca 493
DB	422	cttctgcgaacctatgctggcgccctccgtgcgggggcgcgtcaagctcaaccaacggcca 481
QY	494	gaaggacatcgccgtgaattggcgctggggcctcaccaatgcaagaagtccgagggcatc 553
DB	482	---tgacatcgcaactaactgtgcggggcctgcacacagcgaagaagtgcgagggctc 538
QY	554	tggctctgttacgtcaatgatatcgtcttggcactcctgtgaactgtctaagtatacca 613
DB	539	gggcttctgtctacgtcaatgacatcgtgtctgcacatactcgagctgtcgaagcatcacga 598

QY	614	gagggtcgttactattgacattgatatattcaacatgtgtgacggcgttggaaaggccttcta 673
DB	599	gagagttctgtatgtcgcatalcgcatalcccaacatgtgtgagtgaggagccttcta 658
QY	674	caccacggaccgggtcatgactgtgtccctttcataagtatggagagacttcccaaggaa 733
DB	659	cacaacagatagggttatgactgtgtcgttccacaagtttgttgattatttcccaaggaa 718
QY	734	tgggacctaagcgataccggcgctggcgaagaacaagtattatgtctgtaactaccgcgt 793
DB	719	agggatataccgtgcacatttgggcactcaaaagggaagtactactccctgaatgtccctc 778
QY	794	ccgagacggatgtatgacagagtcctatgaggccaatttcaagccggtcatgtccaagt 853
DB	779	agatgattggatgtatgaaagctaccagtccttllaagccaatcatgycgaagt 838
QY	854	aatgagatgttccagcctagtgcggtgtcttacagtgltgtcagactccctatctgg 913
DB	839	tatgagagtttccgccctgtgtgcagltgtgtcattcagtggtgtgcgattccctgtctgg 898
QY	914	ggatcggttaggtgttcttcaatctatcttcaaaaggacacggccaagtgtgtgaattgt 973
DB	899	ggataggttgggtgtcttcaacctctcaatcaaaagtcacgtgcggatgtgttagtata 958
QY	974	caagagctttaacctgcctatgtctgatgtcgggagcggtgtgttacaccattcgtaagt 1033
DB	959	gaggtcttcaacgcttccattgttgcctctgtgtgtgtgtgatatacaccataagaatgt 1018
QY	1034	tgccggtgtgacatatgagacagctgtggccctggatagacggagatccctaagt 1093
DB	1019	tgcacgtgtgtgttatgagactgtgagctgtccttggccaagagcctgaagacaagt 1078
QY	1094	tcatacaatgactactttgaatacttggaccagatttcaagctccacatcagtcctc 1153
DB	1079	gcctgttaatgactactatgaataacttcggtccagattacactcttcatgttgcaccaag 1138
QY	1154	caatatgactaacgaacaacgaatgacttctggagaagatcaaacagcgaactgtttga 1213
DB	1139	taacatgagagaacaaaatacacgacacaactgtgatatacgtactaaacttctgga 1198
QY	1214	gaacctgaatgtgtccgcgacgcaccttgggggtccaacgcgagcgga---ttcctgagga 1270
DB	1199	taatcttcaaaacttcgacatgtctcctagtgctccactttcaagagagagttcctgacac 1258
QY	1271	cgccatccctgagggaggtggtgcatgagagca 1303
DB	1259	agaataacctggaagaatgaagaatcaagaatga 1291
RESULT 13		
AAX90839		
ID	AAX90839 standard; DNA; 2019 BP.	
XX		
AC	AAX90839;	
XX		
DT	13-JAN-2000 (first entry)	
XX		
DE	Maize histone deacetylase-3 DNA.	
XX		
KW	Maize histone deacetylase; family 1, ZmHD1; promoter regulator; promoter;	
KW	RNA polymerase II; transcription; plant transformation; heterochromatin;	
KW	disease resistance; chromatin assembly; gene activity; toxin screening;	
KW	pathogenicity; disease response promoter; ds.	
XX		
OS	Zea mays.	
XX		
EH	Key	Location/Qualifiers
FT	CDS	140..1459
FT		/*tag= a
FT		/product= "Maize histone deacetylase"
XX		WO9951731-A2.
PN		

XX 14-OCT-1999.
 PD
 XX 02-APR-1999; 99WO-US07370.
 PF
 XX 03-APR-1998; 98US-0080563.
 PR
 XX
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 PI
 PI Baldwin DA, Briggs SP, Crane VC;
 XX
 DR WPI, 1999-611038/52.
 DR P-PSDB; AAY28799.
 XX
 PT
 PT New deacetylase genes, used for producing transgenic plants which have
 PT increased disease resistance -
 XX
 PS
 PS Claim 1; Page 55-58; 87pp; English.
 CC
 CC The present sequence encodes a maize histone deacetylase. This DNA
 CC belongs to family I, ZmHD1 and appears to be regulator of promoters for
 CC RNA polymerase II, for transcription of genes. The nucleotide sequence
 CC can be used to transform plants and increase disease resistance by
 CC altering the levels of heterochromatin, altering chromatin assembly, and
 CC gene activity of the transformed plants. Additionally, compositions find
 CC use in screening for toxins that affect pathogenicity and in determining
 CC which disease response promoters are regulated by histone deacetylase.
 SQ Sequence 2019 BP; 510 A; 532 C; 528 G; 449 T; 0 other;

Query Match	31.1%;	Score 501.8;	DB 20;	Length 2019;
Best Local Similarity	65.38;	Pred. No. 7.3e-142;		
Matches 753; Conservative	0;	Mismatches 397;	Indels 3;	Gaps 1;

QY	121	cgcaggccaccgaggaagaagtctgtactactacgacggatgttggaaattactatt	180
Db	189	cggacgggtcgaagcgccgctctgtacttctaagacgcgaggtggcaactact	248
QY	181	atgcacaagggccaaccaatgaagtcacccaagatccgatgactcataatttgctca	240
Db	249	acgggcagggccaccgccgatgtgaagccgacccgcatccgatgaccacgcgctcgcc	308
QY	241	actatgctctctacccgaaaatgtgaatctatacgccctcacaaagccaatgtgagga	300
Db	309	gtacagcctctctcgaccagatgcaagtgttcgcctcaccctgcccgcgaccgcgacc	368
QY	301	tgaaccaagtaccacagcgatgactcataaattcttgcgtccatccgtccagataaca	360
Db	369	tctgcgcttccaagcgcagcattaagttctctctcctccgctccgtacccccgaaaagc	428
QY	361	tgtcggagtagcagcaagcagatgcagagattcaacglttgtgtaggactgtccagtatcg	420
Db	429	agcagcaccagatccgcgcgtcaagcgcttcaaagctcgcgagagactgcccgcgttcg	488
QY	421	atgcgctgtttgagttctgttcagttgttctactgtgtgttctgttgcgaagtgtctgaaac	480
Db	489	acggtctctacagtttctgttcagacgltacgcggggggtctgttggcggcgccgtcaagc	548
QY	481	ttaataagcagcagacgcagacatcgccgtgaattggcttggggcctgcaccatgcaaga	540
Db	549	tcaaccatggcca---tgatatcgccatcaacttgggcggcgactccaaccagccaaga	605
QY	541	agtcgcagggcatctggtctctgttacgccaatgatalcgtcttggccatccctgaaactgc	600
Db	606	agtcgagggcctccgggttttgctatgttaatgacattgttcctcgccatccctcgagctcc	665
QY	601	taaagtataccaccagaggtgtgtacatltgacattgatataccaacatgtgtgacgcgctgg	660
Db	666	tcaagtaccaccacagcgcttctgtacgttgacattgatataccaacaggggacgcgctgg	725
QY	661	aagagcgcttctaacaccacgacccggtcatgactgtgtcctttcataagtatgagagt	720

Db	726	aggagggcttlltataccaacagaccgggtgatgacagcttcattccaagaagtttggagatt	785
QY	721	acttcccaggaactgsgggaacctacggygataaccggggcttggcaagacaagtattatgtctg	780
Db	786	atttcccttggacacaggggacattcgttgatgttgggcactcaaaaggytaatatattctccc	845
QY	781	ttaactaccgcgtccggaacgggatttgatgacgagcttcataagggcattttcaaccgg	840
Db	846	tgatgttcccttggacgagctgtattgatgatgagagctacaccagtcgttgttcaagccaa	905
QY	841	tcatgtccaagaatagtgagatgttccagcctaagtgtgggtgtcttacagttgtgctcag	900
Db	906	taatgggcaagtgatgtgaggtcttcaaccctgtgtgaagtcgtctccaagtgtgtgcgg	965
QY	901	acttccatcttggggatcgggttaagtttgtctcaatctatctatccaagacacggccaagt	960
Db	966	attcattgtcgggtgacaggtttgggtctgttcaacccttctattaagggtcaccgcagaat	1025
QY	961	gtgtgggaatttgtcaagaagccttaacctgtccatgtcgtatgtctggagggcgttgttaca	1020
Db	1026	gtgtgaagatttatgaggtctcttcaacgctcccgctgtgtcgtctgtgtgtgtgtata	1085
QY	1021	ccattcgtlaacgttgcggcggtcgtgacatatgagacagcgttggccctgtgatacggaga	1080
Db	1086	ccataagaacgctgtgcacggtgtgtgtgctacgagacaggaagttgccctgtgtcatgagc	1145
QY	1081	tcacctaatgagcttccatacaaatgactactttgaatacttttggaccagatttcaagctcc	1140
Db	1146	tcactgacacaagatgtccaccctatgagtactatgagtatttggccagatttaccacttac	1205
QY	1141	acatcagtccttccaatatgtacttaaccagacacagaaatgagtaccttgagagaagatccaac	1200
Db	1206	atgttcgtctccaagtaacatagtgaagataaaacacacacggtcatcaattggtatgcacataaat	1265
QY	1201	agcgactgttttgaagaaccttagaatgtctccgcacgcgaaccttgggttccaacgcgacggcga	1260
Db	1266	caaaactctagataatctttcaaaaactcgcacatgtctcctagtgttcaagtttcaagagc	1325
QY	1261	ttcctgaaggagcg 1273	
Db	1326	gacctcctgagcg 1338	

RESULT	14	
AAZ58259		
ID	AAZ58259	standard; cDNA; 1990 BP.
XX		
AC	AAZ58259;	
XX		
DT	08-MAY-2000	(first entry)
XX		
DE	Rice	histone deacetylase gene 1 (HD1) cDNA contig.
XX		
KW	Chromatin associated protein;	histone deacetylase gene 1; HD1;
KW	rice; transgenic plant;	transcription regulation; ss.
XX		
OS	Oryza sativa.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	3..1484
FT		/*tag= a
FT		/partial
XX		
PN	WO200004177-A1.	
XX		
PD	27-JAN-2000.	
XX		
PF	13-JUL-1999;	99WO-US15807.
XX		
PR	14-JUL-1998;	98US-0092841.
XX		
PA	(DUPO) DU PONT DE NEMOURS & CO E I.	
XX		

(DUPO) DU PONT DE NEMOURS & CO E I.

DR WPI: 2001-258457/27.
DR P-PSDB; AAB67811.
XX
PT Methods for regulating gene expression in transgenic plants, e.g.
PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT elongation), comprises introducing genes encoding histone deacetylase
PT
XX
PS Claim 9; Fig 1A; 91pp; English.
XX
CC The present sequence encodes an Arabidopsis thaliana histone deacetylase
CC designated AtRPD3A. The protein is homologous to yeast RPD3 and HDAC1.
CC The polynucleotide sequence is used in the method of the invention.
CC The specification describes a method for regulating gene expression in
CC transgenic plants. The method comprises modifying histones by introducing
CC chimeric nucleotide sequences which have regulatory elements in operative
CC association with a gene of interest or with a nucleotide sequence
CC encoding histone deacetylase. The method is useful for regulating the
CC developmental, physiological or biochemical pathway within a plant,
CC particularly for repressing ethylene-responsive phenotypes
CC (e.g. inhibition of hypocotyl elongation). The method is also useful
CC as a functional test for identifying a phenotype associated with
CC perturbing a gene. The histone deacetylase genes are useful for
CC altering the development of an organism.
CC
XX
SQ Sequence 1807 BP; 472 A; 386 C; 464 G; 485 T; 0 other;

Query Match 30.98; Score 498.2; DB 22; Length 1807;
Best Local Similarity 64.08; Pred. No. 8.6e-141;
Matches 752; Conservative 0; Mismatches 423; Indels 0; Gaps 0;

QY 134 gagaaagtctgttactacgacgggagtggtgaattractatattgacaagcca 193
Db 189 gagaaagtgttatttactacgacctgagtgctgcgaattactactatgccaaggtca 248
QY 194 cccaatgaagcctcacccgaatccgatgactcataattgtcgtccaactatgtctcta 253
Db 249 tcccatgaagcccatcgcatcgcatgaccatgcccctcctcgtcactaaggtctcct 308
QY 254 ccgaaaaatggaatctatcgccctcacaaagccaatgctgagagatgaccaagtacca 313
Db 309 tcagcatatgcaggttctcaagcccttcctgcccgcgaagctgactctctgcccgtcca 368
QY 314 cagcgatgactacattaaattctgtcgtccatcgtccaagataacatgtcggagtaacag 373
Db 369 cgccgaagcactatgtctcttcttcgcgagcatatccccctgaaagcccaagatcagat 428
QY 374 caagcagatgcagagattcaacgtgtgtaggactgtccagattcgatgacctgttga 433
Db 429 tcgccaacttaagcgcttcaagctgtgtagaagactgtccgctctttagcgccttattc 488
QY 434 gtctgtcagtgtgtctactgtgtgtctgtggaagtgtgtgaaacttaataagcagca 493
Db 489 ctcttgcagacctatgtctgtaggatctgtgtgtgtgtctgtcaagcttaaccacgacct 548
QY 494 gacggacatgcgcgtgaattgggctggggcctgcaccatgcaagaagtccgagcctc 553
Db 549 ctgcgatatgtccatcaactggtgtgtgtctcctcatcacgttaagaagtgcgagcctc 608
QY 554 tggcttctgttaagtcataatgatactgtcttgccatcctgtaactgttaagtataacca 613
Db 609 tggcttctgttaagtcataatgatactgtcttgccatcctgtaactgttaagtataacca 668
QY 614 gaggtgtctgtacattgacattgataattcaccatggtgacgctgtggaagacctcta 673
Db 669 gcgtgtctcttatgtcgatatgatalccaccacgggagtgagtgagagcatttta 728
QY 674 caccacgacgcgggtcatgactgtgtcctttcataaglatgagagtlactccaggaac 733
Db 729 tgtactgacagaggttatgactgtctcgtttcataaattgtgattacttcccggtac 788
QY 734 tggggacctacgggataccggggctcgcaagaagacagtattatgtttaactaccgct 793

Db 789 aggtacattcagatatagttatgttagcggaagtaactatctcctcaatgtaccact. 848
QY 794 ccgagacgggatgtatgacgagttcctatgagggcattttcaagccggtcatgtccaagt 853
Db 849 ggtatgtaagatcgtatgatgagagctatcatctgttattcaagcccatcatgtgggaagt 908
QY 854 aatgagatgttccagcctagtgcgtgtgtcttacagtgtgtcagactcccatctgtg 913
Db 909 tatgaaattttccgaccaggggtgtgtgtattgcaatgtgtgtgactcccatctgtg 968
QY 914 ggatcggttaggtgtgtctcaatcattatatacaaaaggacacgccaagtgtgtgaattgt 973
Db 969 ggatcggttaggtgtgtctcaatcattatatacaaaaggatcagtgtgtgtcaatttat 1028
QY 974 caagagctttaacctgtcctatgtgatgtgtaggagcggtgttacaccattcgttaagt 1033
Db 1029 gagatcgttcaatgttccctactcgtctgtgtgtgtgtgtgtgttacactatccgcaatgt 1088
QY 1034 tgcgcgtgtgacatatgagacagctgtggccctgtgatacggagatccctaagtgt 1093
Db 1089 tgcgcgtgtgtgtgtacgagagctgagtgctgactgtgagttgaagtgaagacaagt 1148
QY 1094 tccatacaatgactacttttgaatactttggaccagatttcaagcttccacatcagtccttc 1153
Db 1149 gccggagcatgaatatattatgaatactttgttccagactatacaacttcaagttgtccaag 1208
QY 1154 caatagacttaaccagaaacggaatgagttaccttgagagaatccaacagcagctgttga 1213
Db 1209 taacatggaataaataagaattctcgtcagatgtcttgaagagatcgcgaatgaccttcca 1268
QY 1214 gaaccttaagaatgtgcgcgacgcaacctgggtccaaacgcaagcgattcctgaggaagc 1273
Db 1269 caatcctctaagcttcaagcatgtctccaagtgtaaccatttcaaggaaagaccacctgatac 1328
QY 1274 catccctgaggaagagtggtgcgatgaggaagcaagagc 1308
Db 1329 agagactcccgaggtgtgatgaaagaccagaagatg 1363

Search completed: June 8, 2002, 12:35:42
Job time: 10479 sec

THIS PAGE BLANK (USPTO)